



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Testing gravity using galaxy-galaxy lensing and clustering amplitudes in KiDS-1000, BOSS and 2dFLenS

Citation for published version:

Blake, C, Amon, A, Asgari, M, Bilicki, M, Dvornik, A, Erben, T, Giblin, B, Glazebrook, K, Heymans, C, Hildebrandt, H, Joachimi, B, Joudaki, S, Kannawadi, A, Kuijken, K, Lidman, C, Parkinson, D, Shan, H, Tröster, T, Busch, JLVD, Wolf, C & Wright, AH 2020, 'Testing gravity using galaxy-galaxy lensing and clustering amplitudes in KiDS-1000, BOSS and 2dFLenS', *Astronomy and Astrophysics*, vol. 642, A158. <https://doi.org/10.1051/0004-6361/202038505>

Digital Object Identifier (DOI):

[10.1051/0004-6361/202038505](https://doi.org/10.1051/0004-6361/202038505)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Astronomy and Astrophysics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Testing gravity using galaxy-galaxy lensing and clustering amplitudes in KiDS-1000, BOSS and 2dFlenS

Chris Blake^{1*}, Alexandra Amon², Marika Asgari³, Maciej Bilicki⁴, Andrej Dvornik⁵, Thomas Erben⁶, Benjamin Giblin³, Karl Glazebrook¹, Catherine Heymans^{3,5}, Hendrik Hildebrandt⁵, Benjamin Joachimi⁷, Shahab Joudaki⁸, Arun Kannawadi^{9,10}, Konrad Kuijken¹⁰, Chris Lidman¹¹, David Parkinson¹², HuanYuan Shan^{13,14}, Tilman Tröster³, Jan Luca van den Busch⁵, Christian Wolf¹¹, and Angus H. Wright⁵

¹ Centre for Astrophysics & Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, VIC 3122, Australia

² Kavli Institute for Particle Astrophysics & Cosmology, P.O. Box 2450, Stanford University, Stanford, CA 94305, USA

³ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK

⁴ Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668, Warsaw, Poland

⁵ Ruhr-University Bochum, Astronomical Institute, German Centre for Cosmological Lensing, Universitätsstr. 150, 44801 Bochum, Germany

⁶ Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany

⁷ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

⁸ Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

⁹ Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA

¹⁰ Leiden Observatory, Leiden University, P.O.Box 9513, 2300RA Leiden, The Netherlands

¹¹ Research School of Astronomy and Astrophysics, Australian National University, Canberra ACT 2611, Australia

¹² Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea

¹³ Shanghai Astronomical Observatory (SHAO), Nandan Road 80, Shanghai 200030, China

¹⁴ University of Chinese Academy of Sciences, Beijing 100049, China

Received date / Accepted date

ABSTRACT

The physics of gravity on cosmological scales affects both the rate of assembly of large-scale structure and the gravitational lensing of background light through this cosmic web. By comparing the amplitude of these different observational signatures, we can construct tests that can distinguish general relativity from its potential modifications. We used the latest weak gravitational lensing dataset from the Kilo-Degree Survey, KiDS-1000, in conjunction with overlapping galaxy spectroscopic redshift surveys, BOSS and 2dFlenS, to perform the most precise existing amplitude-ratio test. We measured the associated E_G statistic with 15 – 20% errors in five $\Delta z = 0.1$ tomographic redshift bins in the range $0.2 < z < 0.7$ on projected scales up to $100 h^{-1}$ Mpc. The scale-independence and redshift-dependence of these measurements are consistent with the theoretical expectation of general relativity in a Universe with matter density $\Omega_m = 0.27 \pm 0.04$. We demonstrate that our results are robust against different analysis choices, including schemes for correcting the effects of source photometric redshift errors, and we compare the performance of angular and projected galaxy-galaxy lensing statistics.

Key words. dark energy – large-scale structure of Universe – gravitational lensing: weak – surveys

1. Introduction

A central goal of modern cosmology is to discover whether the dark energy that appears to fill the Universe is associated with its matter-energy content, laws of gravity, or some alternative physics. A compelling means of distinguishing between these scenarios is to analyse the different observational signatures that are present in the clumpy, inhomogeneous Universe, which powerfully complements measurements of the expansion history of the smooth, homogeneous Universe (e.g. Linder 2005; Wang 2008; Guzzo et al. 2008; Weinberg et al. 2013; Huterer et al. 2015).

Two important observational probes of the inhomogeneous Universe are the peculiar velocities induced in galaxies by the gravitational collapse of large-scale structure, which are statistically imprinted in galaxy redshift surveys as redshift-space dis-

tortions (e.g. Hamilton 1998; Scoccimarro 2004; Song & Percival 2009), and the gravitational lensing of light by the cosmic web, which may be measured using cosmic shear surveys (e.g. Bartelmann & Schneider 2001; Kilbinger 2015; Mandelbaum 2018). These probes are complementary because they allow for the differentiation between the two space-time metric potentials which govern the motion of non-relativistic particles, such as galaxy tracers, and the gravitational deflection of light. The difference or ‘gravitational slip’ between these potentials is predicted to be zero in general relativity, but it may be significant in modified gravity scenarios (e.g. Uzan & Bernardeau 2001; Zhang et al. 2007; Jain & Khoury 2010; Bertschinger 2011; Clifton et al. 2012).

Recent advances in weak gravitational lensing datasets, including the Kilo-Degree Survey (KiDS, Hildebrandt et al. 2020), the Dark Energy Survey (DES, Abbott et al. 2018) and the Subaru Hyper Suprime-Cam Survey (HSC, Hikage et al. 2019), have

* E-mail: cblake@swin.edu.au

led to dramatic improvements in the quality of these observational tests. Gravitational lensing now permits the accurate determination of, and combinations of, important cosmological parameters such as the matter density of the Universe and normalisation of the matter power spectrum, and thereby detailed comparisons with other cosmological probes such as galaxy clustering (Alam et al. 2017a) and the cosmic microwave background radiation (Planck Collaboration et al. 2018). Some of these comparisons have yielded intriguing evidence of ‘tension’ on both small and large scales (e.g. Joudaki et al. 2017; Leauthaud et al. 2017; Lange et al. 2019; Hildebrandt et al. 2020; Asgari et al. 2020b), which are currently unresolved.

In this paper we perform a new study regarding this question using the latest weak gravitational lensing dataset from the Kilo-Degree Survey, KiDS-1000 (Kuijken et al. 2019), in conjunction with overlapping galaxy spectroscopic redshift survey data from the Baryon Oscillation Spectroscopic Survey (BOSS, Reid et al. 2016) and the 2-degree Field Lensing Survey (2dFLenS, Blake et al. 2016a). In particular, we focus on a simple implementation of the lensing-clustering test which compares the amplitude of gravitational lensing around foreground galaxies (commonly known as galaxy-galaxy lensing), tracing low-redshift overdensities, with the amplitude of galaxy velocities induced by these overdensities and measured by redshift-space distortions, which constitutes an amplitude-ratio test. This diagnostic was first proposed by Zhang et al. (2007) as the E_G statistic and implemented in its current form by Reyes et al. (2010) using data from the Sloan Digital Sky Survey. These measurements have subsequently been refined by a series of studies (Blake et al. 2016b; Pullen et al. 2016; Alam et al. 2017b; de la Torre et al. 2017; Amon et al. 2018; Singh et al. 2019; Jullo et al. 2019) which have used new datasets to increase the accuracy of the amplitude-ratio determination, albeit showing some evidence of internal disagreement.

The availability of the KiDS-1000 dataset and associated calibration samples has allowed us to perform the most accurate existing amplitude-ratio test, on projected scales up to $100 h^{-1}$ Mpc, including rigorous systematic-error control. As part of this analysis, we use these datasets and representative simulations to study the efficacy of different corrections for the effects of source photometric redshift errors, comparing different galaxy-galaxy lensing estimators and the relative performance of angular and projected statistics. Our analysis sets the stage for future percent level implementations of these tests using new datasets from the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2016), the 4-metre Multi-Object Spectrograph Telescope (4MOST, de Jong et al. 2019), the Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2019), and the *Euclid* satellite (Laureijs et al. 2011).

This paper is structured as follows: in Sect. 2 we review the theoretical correlations between weak lensing and overdensity observables, on which galaxy-galaxy lensing studies are based. In Sect. 3 we summarise the angular and projected galaxy-galaxy lensing estimators derived from these correlations, with particular attention to the effect of source photometric redshift errors. In Sect. 4 we introduce the amplitude-ratio test between galaxy-galaxy lensing and clustering observables, constructed from annular differential surface density statistics, and in Sect. 5 we derive the analytical covariances of these estimators in the Gaussian approximation, including the effects of the survey window function. We introduce the KiDS-1000 weak lensing and overlapping Luminous Red Galaxy (LRG) spectroscopic datasets in Sect. 6. We create representative survey mock catalogues in Sect. 7, which we use to verify our cosmological analysis in Sect. 8.

Finally, we describe the results of our cosmological tests applied to the KiDS-LRG datasets in Sect. 9. We summarise our investigation in Sect. 10.

2. Theory

In this section we briefly review the theoretical expressions for the auto- and cross-correlations between weak gravitational lensing and galaxy overdensity observables, which form the basis of galaxy-galaxy lensing studies.

2.1. Lensing convergence and tangential shear

The observable effects of weak gravitational lensing, on a source located at co-moving co-ordinate χ_s in sky direction $\hat{\Omega}$, can be expressed in terms of the lensing convergence κ (for reviews, see Bartelmann & Schneider 2001; Kilbinger 2015; Mandelbaum 2018). The convergence is a weighted integral over co-moving distance χ of the matter overdensity δ_m along the line-of-sight, which we can write as

$$\kappa(\chi_s, \hat{\Omega}) = \frac{3\Omega_m H_0^2}{2c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s} \frac{\delta_m(\chi, \hat{\Omega})}{a(\chi)}, \quad (1)$$

assuming (throughout this paper) a spatially-flat Universe, where Ω_m is the matter density as a fraction of the critical density, H_0 is the Hubble parameter, c is the speed of light, and $a = 1/(1+z)$ is the cosmic scale factor at redshift z . We can conveniently write Eq. 1 in terms of the critical surface mass density at a lens plane at co-moving distance χ_1 ,

$$\Sigma_c(\chi_1, \chi_s) = \frac{c^2}{4\pi G} \frac{\chi_s}{(1+z_1)\chi_1(\chi_s - \chi_1)}, \quad (2)$$

where G is the gravitational constant, and $\chi_s > \chi_1$. Hence,

$$\kappa(\chi_s, \hat{\Omega}) = \bar{\rho}_m \int_0^{\chi_s} d\chi \Sigma_c^{-1}(\chi, \chi_s) \delta_m(\chi, \hat{\Omega}), \quad (3)$$

where $\bar{\rho}_m$ is the mean matter density.¹

Suppose that the overdensity is associated with an isolated lens galaxy at distance χ_1 in an otherwise homogeneous Universe. In this case, Eq. 3 may be written in the form

$$\kappa(\chi_1, \chi_s, \hat{\Omega}) \approx \bar{\rho}_m \Sigma_c^{-1}(\chi_1, \chi_s) \int_0^{\chi_s} d\chi \delta_m(\chi, \hat{\Omega}). \quad (4)$$

Eq. 4 motivates that the weak lensing observable can be related to the projected mass density around the lens, $\Sigma = \int \rho_m d\chi$, where $\delta_m = \rho_m/\bar{\rho}_m - 1$. The convergence may be written in terms of this quantity as

$$\kappa(\chi_1, \chi_s, \hat{\Omega}) \approx \Sigma_c^{-1}(\chi_1, \chi_s) (\Sigma - \bar{\Sigma}), \quad (5)$$

where $\bar{\Sigma} = \int \bar{\rho} d\chi$ represents the average background, emphasising that gravitational lensing traces the increment between the mass density and the background.

The average tangential shear γ_t at angular separation θ from an axisymmetric lens is related to the convergence as

$$\langle \gamma_t(\theta) \rangle = \langle \bar{\kappa}(<\theta) \rangle - \langle \kappa(\theta) \rangle, \quad (6)$$

¹ We refer the reader to Dvornik et al. (2018) Appendix C for a full discussion of the different definitions of Σ_c that have been adopted in the literature.

where $\bar{\kappa}(< \theta)$ is the mean convergence within separation θ . At the location of the lens, angular separations are related to projected separations as $R = \chi(z_l) \theta$. Defining the differential projected surface mass density around the lens as a function of projected separation,

$$\Delta\Sigma(R) = \bar{\Sigma}(< R) - \Sigma(R), \quad (7)$$

where,

$$\bar{\Sigma}(< R) = \frac{2}{R^2} \int_0^R R' \Sigma(R') dR', \quad (8)$$

we find that for a single source-lens pair at distances χ_1 and χ_s (omitting the angled brackets),

$$\gamma_t(\theta) = \Sigma_c^{-1}(\chi_1, \chi_s) \Delta\Sigma(R). \quad (9)$$

2.2. Galaxy-convergence cross-correlation

For an ensemble of sources with distance probability distribution $p_s(\chi)$ (normalised such that $\int p_s(\chi) d\chi = 1$), the total convergence in a given sky direction is

$$\begin{aligned} \kappa(\hat{\Omega}) &= \int d\chi_s p_s(\chi_s) \kappa(\chi_s, \hat{\Omega}) \\ &= \bar{\rho}_m \int_0^\infty d\chi \bar{\Sigma}_c^{-1}(\chi) \delta_m(\chi, \hat{\Omega}), \end{aligned} \quad (10)$$

where,

$$\bar{\Sigma}_c^{-1}(\chi) = \int_\chi^\infty d\chi_s p_s(\chi_s) \Sigma_c^{-1}(\chi, \chi_s), \quad (11)$$

with the lower limit of the integral applying because $\Sigma_c^{-1}(\chi_1, \chi_s) = 0$ for $\chi_s < \chi_1$. We consider forming the angular cross-correlation function of this convergence field with the projected number overdensity of an ensemble of lenses with distance probability distribution $p_l(\chi)$,

$$\delta_{g,2D}(\hat{\Omega}) = \int d\chi p_l(\chi) \delta_g(\chi, \hat{\Omega}). \quad (12)$$

The galaxy-convergence cross-correlation function at angular separation θ is,

$$\omega_{g\kappa}(\theta) = \langle \kappa(\hat{\Omega}) \delta_{g,2D}(\hat{\Omega} + \theta) \rangle. \quad (13)$$

Expressing the overdensity fields in terms of their Fourier components we find, after some algebra,

$$\omega_{g\kappa}(\theta) = \int \frac{d^2\ell}{(2\pi)^2} C_{g\kappa}(\ell) e^{-i\ell \cdot \theta}, \quad (14)$$

where ℓ is a 2D Fourier wavevector, and the corresponding angular cross-power spectrum $C_{g\kappa}(\ell)$ is given by (Guzik & Seljak 2001; Hu & Jain 2004; Joachimi & Bridle 2010),

$$C_{g\kappa}(\ell) = \bar{\rho}_m \int d\chi p_l(\chi) \frac{\bar{\Sigma}_c^{-1}(\chi)}{\chi^2} P_{gm}\left(\frac{\ell}{\chi}, \chi\right), \quad (15)$$

where $P_{gm}(k, \chi)$ is the 3D galaxy-matter cross-power spectrum at wavenumber k and distance χ . Taking the azimuthal average of Eq. 14 over all directions θ , the complex exponential integrates to a Bessel function of the first kind, $J_0(x)$, such that,

$$\omega_{g\kappa}(\theta) = \int \frac{d\ell \ell}{(2\pi)^2} C_{g\kappa}(\ell) J_0(\ell\theta) = \int \frac{d\ell \ell}{2\pi} C_{g\kappa}(\ell) J_0(\ell\theta). \quad (16)$$

Using Eq. 6 and Bessel function identities, we can then obtain an expression for the statistical average tangential shear around an ensemble of lenses,

$$\gamma_t(\theta) = \overline{\omega_{g\kappa}}(< \theta) - \omega_{g\kappa}(\theta) = \int \frac{d\ell \ell}{2\pi} C_{g\kappa}(\ell) J_2(\ell\theta). \quad (17)$$

Likewise, we can generalise Eq. 9 to apply to broad source and lens distributions:

$$\gamma_t(\theta) = \int d\chi p_l(\chi) \bar{\Sigma}_c^{-1}(\chi) \Delta\Sigma(R, \chi). \quad (18)$$

Comparing the formulations of Eqs. 17 and 18 allows us to demonstrate that,

$$\Sigma(R) = \bar{\rho}_m \int_{-\infty}^{\infty} d\Pi \left[1 + \xi_{gm}(R, \Pi) \right], \quad (19)$$

in terms of the 3D galaxy-matter cross-correlation function $\xi_{gm}(R, \Pi)$ at projected separation R and line-of-sight separation Π , where the constant term ‘1+’ cancels out in the evaluation of the observable $\Delta\Sigma$. After some algebra we find,

$$\Delta\Sigma(R) = \bar{\rho}_m \int_0^\infty dr W(r, R) \xi_{gm}(r), \quad (20)$$

where,

$$W(r, R) = \frac{4r^2}{R^2} - \left[\frac{4r \sqrt{r^2 - R^2}}{R^2} + \frac{2r}{\sqrt{r^2 - R^2}} \right] H(r - R), \quad (21)$$

where $H(x) = 0$ if $x < 0$ and $H(x) = 1$ if $x > 0$ is the Heaviside step function. The relations in this section make the approximations of using the Limber equation (Limber 1953) and neglecting additional effects such as cosmic magnification (Unruh et al. 2019) and intrinsic alignments (Joachimi et al. 2015).

2.3. Auto-correlation functions

In order to determine the analytical covariance in Sect. 5, we also need expressions for the auto-correlation functions of the convergence, $\omega_{\kappa\kappa}(\theta) = \int \frac{d^2\ell}{(2\pi)^2} C_{\kappa\kappa}(\ell) e^{-i\ell \cdot \theta}$, and the galaxy overdensity, $\omega_{gg}(\theta) = \int \frac{d^2\ell}{(2\pi)^2} C_{gg}(\ell) e^{-i\ell \cdot \theta}$. Given two source populations with distance probability distributions $p_{s,1}(\chi)$ and $p_{s,2}(\chi)$, and associated integrated critical density functions $\bar{\Sigma}_{c,1}^{-1}$ and $\bar{\Sigma}_{c,2}^{-1}$, the angular power spectrum of the convergence is given by,

$$C_{\kappa\kappa}(\ell) = \bar{\rho}_m^2 \int d\chi \frac{\bar{\Sigma}_{c,1}^{-1}(\chi) \bar{\Sigma}_{c,2}^{-1}(\chi)}{\chi^2} P_{mm}\left(\frac{\ell}{\chi}, \chi\right), \quad (22)$$

where $P_{mm}(k, \chi)$ is the 3D (non-linear) matter power spectrum at wavenumber k and distance χ . Likewise, for two projected galaxy overdensity fields with distance probability distributions $p_{l,1}(\chi)$ and $p_{l,2}(\chi)$, the angular power spectrum is,

$$C_{gg}(\ell) = \int d\chi \frac{p_{l,1}(\chi) p_{l,2}(\chi)}{\chi^2} P_{gg}\left(\frac{\ell}{\chi}, \chi\right), \quad (23)$$

where $P_{gg}(k, \chi)$ is the 3D galaxy power spectrum.

2.4. Bias model

We computed the linear matter power spectrum $P_L(k)$ in our models using the CAMB software package (Lewis et al. 2000), and evaluated the non-linear matter power spectrum $P_{\text{mm}}(k)$ including the ‘halofit’ corrections (Smith et al. 2003; Takahashi et al. 2012, we define the fiducial cosmological parameters used for the simulation and data analysis in subsequent sections). We adopted a model for the non-linear galaxy-galaxy and galaxy-matter 2-point functions, appearing in Eqs. 15 and 23, following Baldauf et al. (2010) and Mandelbaum et al. (2013). This model assumes a local, non-linear galaxy bias relation via a Taylor expansion of the galaxy density field in terms of the matter overdensity, $\delta_g = b_L \delta_m + \frac{1}{2} b_{\text{NL}} \delta_m^2 + \dots$, defining a linear bias parameter b_L and non-linear bias parameter b_{NL} . The auto- and cross-correlation statistics in this model can be written in the form (McDonald 2006; Smith et al. 2009),

$$\begin{aligned} \xi_{\text{gg}} &= b_L^2 \xi_{\text{mm}} + 2 b_L b_{\text{NL}} \xi_A + \frac{1}{2} b_{\text{NL}}^2 \xi_B, \\ \xi_{\text{gm}} &= b_L \xi_{\text{mm}} + b_{\text{NL}} \xi_A, \end{aligned} \quad (24)$$

where ξ_{mm} is the correlation function corresponding to $P_{\text{mm}}(k)$, and ξ_A and ξ_B are obtained by computing the Fourier transforms of,

$$\begin{aligned} A(k) &= \int \frac{d^3 q}{(2\pi)^3} F_2(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_L(q) P_L(|\mathbf{k} - \mathbf{q}|), \\ B(k) &= \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|), \end{aligned} \quad (25)$$

which depend on the mode-coupling kernel in standard perturbation theory,

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \left(\frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \right)^2. \quad (26)$$

We evaluated these integrals using the FAST software package (McEwen et al. 2016) and note that $\xi_B = \xi_L^2$, where ξ_L is the correlation function corresponding to $P_L(k)$. This model is only expected to be valid on scales exceeding the virial radius of dark matter haloes, since it does not address halo exclusion, the distribution of galaxies within haloes, or other forms of stochastic or non-local effects (Asgari et al. 2020a). However, this 2-parameter bias model is adequate for our large-scale analysis, which we verify using representative mock catalogues in Sect. 8.

3. Estimators

In this section we specify estimators that may be used to measure $\gamma_t(\theta)$ and $\Delta\Sigma(R)$ from ensembles of sources and lenses, and discuss how estimates of $\Delta\Sigma(R)$ are affected by uncertainties in source distances.

3.1. Average tangential shear $\gamma_t(\theta)$

We can estimate the average tangential shear of a set of sources (s) around lenses (l) by evaluating the following expression (Mandelbaum et al. 2006), which also utilises an unclustered random lens catalogue (r) with the same selection function as the lenses:

$$\hat{\gamma}_t(\theta) = \frac{\sum_{\text{ls}} w_l w_s e_{t,\text{ls}} - \sum_{\text{rs}} w_r w_s e_{t,\text{rs}}}{\sum_{\text{rs}} w_r w_s}. \quad (27)$$

The sums in Eq. 27 are taken over pairs of sources and lenses with angular separations within a bin around θ , w_i are weights applied to the different samples (normalised such that $\sum_l w_l = \sum_r w_r$), and e_t indicates the tangential ellipticity of the source, projected onto an axis normal to the line joining the source and lens (or random lens).

Eq. 27 involves the random lens catalogue in two places. First, the tangential shear of sources around random lenses is subtracted from the data signal. The subtracted term has an expectation value of zero, but significantly decreases the variance of the estimator at large separations (Singh et al. 2017). Second, the estimator is normalised by a sum over pairs of sources and random lenses, rather than data lenses. This ensures that the estimator is unbiased: the alternative estimator, $\hat{\gamma}_t = \sum_{\text{ls}} w_l w_s e_{t,\text{ls}} / \sum_{\text{ls}} w_l w_s$, is biased by any source-lens clustering (if the angular cross-correlation function $\omega_{\text{ls}}(\theta) \neq 0$), which would modify the denominator of the expression but not the numerator. The magnitude of this effect is sometimes known as the ‘boost’ factor (Sheldon et al. 2004),

$$B(\theta) = \frac{\sum_{\text{ls}} w_l w_s}{\sum_{\text{rs}} w_r w_s}, \quad (28)$$

where the sums are again taken over source-lens pairs with angular separations within a given bin. We note that $\langle B(\theta) \rangle = 1 + \omega_{\text{ls}}(\theta)$ for unity weights.

3.2. Projected mass density $\Delta\Sigma(R)$

Assuming the source and lens distances are known, each source-lens pair may be used to estimate the projected mass density around the lenses by inverting Eq. 9:

$$\hat{\Delta\Sigma}(R) = e_t(R/\chi_l) \Sigma_c(\chi_l, \chi_s). \quad (29)$$

For an ensemble of sources and lenses, the mean projected mass density may then be estimated by an expression analogous to Eq. 27 (Singh et al. 2017),

$$\begin{aligned} \hat{\Delta\Sigma}(R) &= \frac{\sum_{\text{ls}} w_l w_s w_{\text{ls}} e_{t,\text{ls}}(R/\chi_l) \Sigma_c(\chi_l, \chi_s)}{\sum_{\text{rs}} w_r w_s w_{\text{rs}}} \\ &\quad - \frac{\sum_{\text{rs}} w_r w_s w_{\text{rs}} e_{t,\text{rs}}(R/\chi_r) \Sigma_c(\chi_r, \chi_s)}{\sum_{\text{rs}} w_r w_s w_{\text{rs}}}, \end{aligned} \quad (30)$$

where we have allowed for an additional pair weight between sources and lenses, w_{ls} , and random lenses, w_{rs} . Assuming a constant shape noise in e_t , the noise in the estimate of $\Delta\Sigma(R) = e_t \Sigma_c$ from each source-lens pair is proportional to Σ_c , hence the optimal inverse-variance weight is $w_{\text{ls}} \propto \Sigma_c^{-2}$, and the weighted estimator may be written,

$$\begin{aligned} \hat{\Delta\Sigma}(R) &= \frac{\sum_{\text{ls}} w_l w_s e_{t,\text{ls}}(R/\chi_l) \Sigma_{c,\text{ls}}^{-1}}{\sum_{\text{rs}} w_r w_s \Sigma_{c,\text{rs}}^{-2}} \\ &\quad - \frac{\sum_{\text{rs}} w_r w_s e_{t,\text{rs}}(R/\chi_r) \Sigma_{c,\text{rs}}^{-1}}{\sum_{\text{rs}} w_r w_s \Sigma_{c,\text{rs}}^{-2}}. \end{aligned} \quad (31)$$

3.3. Photo- z dilution correction for $\Delta\Sigma(R)$

The difficulty faced when determining $\Delta\Sigma$ is that source distances are typically only accessible through photometric redshifts and may contain significant errors, leading to a bias in the estimate through incorrect scaling factors Σ_c (we assume in this discussion that spectroscopic lens distances are available). For example, sources may apparently lie behind lenses according to their photometric redshift, whilst in fact being positioned in front of the lenses and contributing no galaxy-galaxy lensing signal, creating a downward bias in the measurement.

For a single source-lens pair, the estimated value of Σ_c for the pair based on the source photometric redshift, $\Sigma_{c,lp}$, may differ from its true value based on the source spectroscopic redshift, $\Sigma_{c,ls}$,

$$\hat{\Delta\Sigma} = e_t \Sigma_{c,lp} = \left(\frac{\Delta\Sigma^{\text{true}}}{\Sigma_{c,ls}} \right) \Sigma_{c,lp} = \left(\frac{\Sigma_{c,lp}}{\Sigma_{c,ls}} \right) \Delta\Sigma^{\text{true}}. \quad (32)$$

Combining many source-lens pairs allowing for a pair weight w_{ls} we find,

$$\hat{\Delta\Sigma} = \frac{\sum_{ls} w_{ls} \left(\frac{\Sigma_{c,lp}}{\Sigma_{c,ls}} \right) \Delta\Sigma^{\text{true}}}{\sum_{ls} w_{ls}}. \quad (33)$$

Using the optimal weight $w_{ls} \propto \Sigma_{c,lp}^{-2}$ this expression may be written,

$$\hat{\Delta\Sigma} = \frac{\sum_{ls} \Sigma_{c,lp}^{-1} \Sigma_{c,ls}^{-1} \Delta\Sigma^{\text{true}}}{\sum_{ls} \Sigma_{c,lp}^{-2}}. \quad (34)$$

The estimated value of $\Delta\Sigma$ hence contains a multiplicative bias, $\Delta\Sigma^{\text{true}} = f_{\text{bias}} \langle \hat{\Delta\Sigma} \rangle$ where,

$$f_{\text{bias}} = \frac{\sum_{ls} \Sigma_{c,lp}^{-2}}{\sum_{ls} \Sigma_{c,lp}^{-1} \Sigma_{c,ls}^{-1}} = \frac{\sum_{ls} w_{ls}}{\sum_{ls} w_{ls} \Sigma_{c,lp} \Sigma_{c,ls}^{-1}}. \quad (35)$$

This multiplicative correction factor may be estimated at each lens redshift from a representative subset of sources with complete spectroscopic and photometric redshift information, by evaluating the sums in the numerator and denominator of Eq. 35 (Nakajima et al. 2012).

An alternative formulation of the photo- z dilution correction may be derived from the statistical distance distribution of the sources. Provided that the lens distribution is sufficiently narrow, Eq. 18 indicates that an unbiased estimate of $\Delta\Sigma$ from each lens-source pair is,

$$\hat{\Delta\Sigma}(R) = e_t(R/\chi_1) \left[\overline{\Sigma_c^{-1}}(\chi_1) \right]^{-1}, \quad (36)$$

where $\overline{\Sigma_c^{-1}}$ is evaluated from Eq. 11 using the source distribution $p_s(\chi)$. This motivates an alternative estimator mirroring Eq. 31 (Sheldon et al. 2004; Miyatake et al. 2015; Blake et al. 2016b),

$$\hat{\Delta\Sigma}(R) = \frac{\sum_{ls} w_l w_s e_{t,ls}(R/\chi_1) \overline{\Sigma_{c,ls}^{-1}}}{\sum_{rs} w_r w_s \left(\overline{\Sigma_{c,rs}^{-1}} \right)^2} - \frac{\sum_{rs} w_r w_s e_{t,rs}(R/\chi_r) \overline{\Sigma_{c,rs}^{-1}}}{\sum_{rs} w_r w_s \left(\overline{\Sigma_{c,rs}^{-1}} \right)^2}. \quad (37)$$

The accuracy of these potential photo- z dilution corrections must be assessed via simulations, which we consider in Sect. 8. We trialled both point-based and distribution-based correction methods in our analysis.

4. Amplitude-ratio test

In this section we construct test statistics which utilise the relative amplitudes of galaxy clustering and galaxy-galaxy lensing to test cosmological models. We first define the input statistics for these tests.

4.1. Projected clustering $w_p(R)$

The amplitude of galaxy-galaxy lensing is sensitive to the distribution of matter around lens galaxies, projected along the line-of-sight. We can obtain an analogous projected quantity for lens galaxy clustering by integrating the 3D galaxy auto-correlation function, ξ_{gg} along the line-of-sight,

$$w_p(R) = \int_{-\infty}^{\infty} d\Pi \xi_{gg}(R, \Pi), \quad (38)$$

where Π is the line-of-sight separation. This formulation has the additional feature of reducing sensitivity of the clustering statistics to redshift-space distortions, which modulate the apparent radial separations Π between galaxy pairs.

We can estimate $w_p(R)$ for a galaxy sample by measuring the galaxy correlation function in (R, Π) separation bins, and summing over the Π direction in the range $0 < \Pi < \Pi_{\text{max}}$:

$$\hat{w}_p(R) = 2 \sum_{\text{bins } i} \Delta\Pi_i \hat{\xi}_{gg}(R, \Pi). \quad (39)$$

4.2. The Upsilon statistics, $\Upsilon_{gm}(R)$ and $\Upsilon_{gg}(R)$

Eq. 7 demonstrates that the amplitude of $\Delta\Sigma(R)$ around lens galaxies depends on the surface density of matter across a range of smaller scales from zero to R , and hence on the galaxy-matter cross-correlation coefficient at these scales. Given that this cross-correlation is a complex function which is difficult to model from first principles, it is beneficial to reduce this sensitivity to small-scale information using the annular differential surface density statistic (Reyes et al. 2010; Baldauf et al. 2010; Mandelbaum et al. 2013),

$$\begin{aligned} \Upsilon_{gm}(R, R_0) &= \Delta\Sigma(R) - \frac{R_0^2}{R^2} \Delta\Sigma(R_0) \\ &= \frac{2}{R^2} \int_{R_0}^R dR' R' \Sigma(R') - \Sigma(R) + \frac{R_0^2}{R^2} \Sigma(R_0), \end{aligned} \quad (40)$$

which is defined such that $\Upsilon_{gm} = 0$ at some small-scale limit $R = R_0$, chosen to be large enough to reduce the main systematic effects (typically, R_0 is somewhat larger than the size scale of dark matter haloes). In this sense, the cumulative effect from the cross-correlation function at scales $R < R_0$ is cancelled, although it is not the case that this small-scale suppression translates to Fourier space (Baldauf et al. 2010; Asgari et al. 2020a; Park et al. 2020). In any case, the efficacy of these statistics and choice of the R_0 value must be validated using simulations, as we consider below.

The corresponding quantity suppressing the small-scale contribution to the galaxy auto-correlations is (Reyes et al. 2010),

$$\Upsilon_{\text{gg}}(R, R_0) = \rho_c \left[\frac{2}{R^2} \int_{R_0}^R dR' R' w_p(R') - w_p(R) + \frac{R_0^2}{R^2} w_p(R_0) \right], \quad (41)$$

where ρ_c is the critical matter density. We note that if w_p is defined as a step-wise function in bins R_i (with bin limits $R_{i,\min}$ and $R_{i,\max}$) then Eq. 41 may be written in the useful form,

$$\Upsilon_{\text{gg}}(R, R_0) = \frac{\rho_c}{R^2} \sum_{i=j}^k C_i w_p(R_i), \quad (42)$$

where (k, j) are the bins containing (R, R_0) , and

$$C_i = \begin{cases} R_{i,\max}^2 & i = j \\ R_{i,\max}^2 - R_{i,\min}^2 & j < i < k \\ -R_{i,\min}^2 & i = k \end{cases} \quad (43)$$

For convenience we chose R_0 to coincide with the centre of a separation bin, such that we could use the direct measurements of $\Delta\Sigma(R_0)$ and $w_p(R_0)$ in Eqs. 40 and 41 without interpolation between bins (we will show below that our results are not sensitive to the choice of R_0).

4.3. The E_G test statistic

The relative amplitudes of weak gravitational lensing and the rate of assembly of large-scale structure depend on the ‘gravitational slip’ or difference between the two space-time metric potentials. This signature is absent in general relativity but may be significant in modified gravity scenarios (Uzan & Bernardeau 2001; Zhang et al. 2007; Jain & Khoury 2010; Bertschinger 2011; Clifton et al. 2012).

Zhang et al. (2007) proposed that these amplitudes might be compared by connecting the velocity field and lensing signal generated by a given set of matter overdensities, probed via redshift-space distortions and galaxy-galaxy lensing, respectively. Reyes et al. (2010) implemented this consistency test by constructing the statistic,

$$E_G(R) = \frac{1}{\beta} \frac{\Upsilon_{\text{gm}}(R, R_0)}{\Upsilon_{\text{gg}}(R, R_0)}, \quad (44)$$

where $\beta = f/b_L$ is the redshift-space distortion parameter which governs the observed dependence of the strength of galaxy clustering on the angle to the line-of-sight, in terms of the linear growth rate of a perturbation, $f = d \ln \delta / d \ln a$. Eq. 44 is independent of the linear galaxy bias b_L and the amplitude of matter clustering σ_8 , given that $\beta \propto 1/b_L$, $\Upsilon_{\text{gm}} \propto b_L \sigma_8^2$ and $\Upsilon_{\text{gg}} \propto b_L^2 \sigma_8^2$. The prediction of linear perturbation theory for general relativity in a Λ CDM Universe is a scale-independent value $E_G(z) = \Omega_m(z=0)/f(z)$, although see Leonard et al. (2015) for a detailed discussion of this approximation.

5. Covariance of estimators

In this section we present analytical formulations in the Gaussian approximation for the covariance of estimates of $\gamma_t(\theta)$ and $\Delta\Sigma(R)$, and model how this covariance is modulated by the presence of a survey mask (that is, by edge effects). Our covariance determination hence neglects non-Gaussian and super-sample variance components. This is a reasonable approximation in the context of the current analysis as these terms are subdominant (we refer the reader to Joachimi et al. (in prep.) for more details on the relative amplitude of the different covariance terms in the context of KiDS-1000).

5.1. Covariance of average tangential shear

In Appendix A we derive the covariance of γ_t averaged within angular bins θ_m and θ_n :

$$\text{Cov}[\gamma_t^{ij}(\theta_m), \gamma_t^{kl}(\theta_n)] = \frac{1}{\Omega} \int \frac{d\ell \ell}{2\pi} \sigma^2(\ell) \overline{J_{2,m}(\ell)} \overline{J_{2,n}(\ell)}, \quad (45)$$

where γ_t^{ij} denotes the average tangential shear of source sample j around lens sample i , Ω is the total survey angular area in steradians, and $\overline{J_{2,n}(\ell)} = \int_{\theta_{1,n}}^{\theta_{2,n}} \frac{2\pi\theta d\theta}{\Omega_n} J_2(\ell\theta)$, where the integral is between the bin limits θ_1 and θ_2 and Ω_n is angular area of bin n (i.e. the area of the annulus between the bin limits). The variance $\sigma^2(\ell)$ is given by the expression for Gaussian random fields (e.g. Hu & Jain 2004; Bernstein 2009; Joachimi & Bridle 2010; Krause & Eifler 2017; Euclid Collaboration et al. 2019),

$$\sigma^2(\ell) = C_{\text{gk}}^{il}(\ell) C_{\text{gk}}^{kj}(\ell) + [C_{\text{kk}}^{jl}(\ell) + N_{\text{kk}}^j \delta_{jl}^{\text{K}}] [C_{\text{gg}}^{ik}(\ell) + N_{\text{gg}}^i \delta_{ik}^{\text{K}}], \quad (46)$$

where δ_{ij}^{K} is the Kronecker delta. The angular auto- and cross-power spectra appearing in Eq. 46 may be evaluated using the expressions in Sect. 2, and the noise terms are $N_{\text{kk}}^i = \sigma_e^2/\bar{n}_s^i$ and $N_{\text{gg}}^i = 1/\bar{n}_l^i$, where σ_e is the shape noise and \bar{n}_l^i and \bar{n}_s^i are the angular lens and source densities of sample i in units of per steradian.

5.2. Covariance of projected mass density

The covariance of $\Delta\Sigma$ may be deduced from the covariance of γ_t using $\Delta\Sigma(R) = \gamma_t(\theta)/\Sigma_c^{-1}$, and by scaling angular separations to projected separations at an effective lens distance χ_1 using $\theta = R/\chi_1$ (Singh et al. 2017; Dvornik et al. 2018; Shirasaki & Takada 2018). We can map multipoles ℓ to the projected wavevector $k = \ell/\chi_1$ such that,

$$\text{Cov}[\Delta\Sigma^{ij}(R), \Delta\Sigma^{kl}(R')] = \frac{1}{\Omega} \int \frac{dk k}{2\pi} \sigma^2(k) \overline{J_2(kR)} \overline{J_2(kR')}, \quad (47)$$

where we now express the variance in terms of projected power spectra,

$$\sigma^2(k) = P_{\text{gk}}^{il}(k) P_{\text{gk}}^{kj}(k) + [P_{\text{kk}}^{jl}(k) + N_{\text{kk}}^j \delta_{jl}^{\text{K}}] [P_{\text{gg}}^{ik}(k) + N_{\text{gg}}^i \delta_{ik}^{\text{K}}]. \quad (48)$$

The power spectra are given by the following relations:

$$P_{\text{gk}}(k) = \frac{\chi_1^2 C_{\text{gk}}(k\chi_1)}{[\Sigma_c^{-1}(\chi_1)]^2} \approx \bar{\rho}_m \int d\chi p_1(\chi) P_{\text{gm}}(k, \chi), \quad (49)$$

$$P_{\text{kk}}(k) = \chi_1^2 C_{\text{kk}}(k\chi_1) = \chi_1^2 \bar{\rho}_m^2 \int d\chi \left[\frac{\Sigma_{c,1}^{-1}(\chi) \Sigma_{c,2}^{-1}(\chi)}{\Sigma_{c,1}^{-1}(\chi_1) \Sigma_{c,2}^{-1}(\chi_1)} \right] \left(\frac{\chi_1^2}{\chi^2} \right) P_{\text{mm}}\left(\frac{k\chi_1}{\chi}, \chi\right), \quad (50)$$

$$P_{\text{gg}}(k) = \chi_1^2 C_{\text{gg}}(k\chi_1) \approx \int d\chi p_1(\chi) p_2(\chi) P_{\text{gg}}(k, \chi), \quad (51)$$

and the noise terms are,

$$N_{\text{kk}} = \frac{\sigma_e^2 \chi_1^2}{\bar{n}_s [\Sigma_c^{-1}(\chi_1)]^2}, \quad N_{\text{gg}} = \frac{\chi_1^2}{\bar{n}_l}. \quad (52)$$

5.3. Covariance of remaining statistics

The expression for the analytical covariance of $w_p(R)$ may be derived as (see also, Singh et al. 2017),

$$\text{Cov}[w_p(R), w_p(R')] = \frac{2L_{\parallel} \Pi_{\max}}{\Omega} \int \frac{dk k}{2\pi} \sigma^2(k) J_0(kR) J_0(kR'), \quad (53)$$

where L_{\parallel} is the total co-moving depth of the lens redshift slice and the expression for the variance is,

$$\sigma^2(k) = [P_{\text{gg}}(k) + N_{\text{gg}}]^2, \quad (54)$$

where $P_{\text{gg}}(k)$ and N_{gg} are the 2D projected power spectra and noise as defined in Sect. 5.2.

We determined the analytical covariance of $\Upsilon_{\text{gm}}(R, R_0)$ from the covariance of $\Delta\Sigma(R)$:

$$\begin{aligned} \text{Cov}[\Upsilon_{\text{gm}}(R, R_0), \Upsilon_{\text{gm}}(R', R_0)] &= \text{Cov}[\Delta\Sigma(R), \Delta\Sigma(R')] \\ &- \frac{R_0^2}{R'^2} \text{Cov}[\Delta\Sigma(R), \Delta\Sigma(R_0)] - \frac{R_0^2}{R^2} \text{Cov}[\Delta\Sigma(R'), \Delta\Sigma(R_0)] \\ &+ \frac{R_0^4}{R^2 R'^2} \text{Var}[\Delta\Sigma(R_0)]. \end{aligned} \quad (55)$$

For the case of $\Upsilon_{\text{gg}}(R, R_0)$, we propagated the covariance using Eq. 42:

$$\begin{aligned} \text{Cov}[\Upsilon_{\text{gg}}(R, R_0), \Upsilon_{\text{gg}}(R', R_0)] &= \\ \frac{\rho_c^2}{R^2 R'^2} \sum_i \sum_j C_i C_j \text{Cov}[w_p(R_i), w_p(R_j)]. \end{aligned} \quad (56)$$

We evaluated the covariance of the E_G statistic, where required, by assuming small fluctuations in the variables in Eq. 44 with respect to their mean, neglecting any correlations between the measurements:

$$\begin{aligned} \frac{\text{Cov}[E_G(R) E_G(R')]}{E_G(R) E_G(R')} &= \frac{\text{Cov}[\Upsilon_{\text{gm}}(R, R_0), \Upsilon_{\text{gm}}(R', R_0)]}{\Upsilon_{\text{gm}}(R, R_0) \Upsilon_{\text{gm}}(R', R_0)} \\ &+ \frac{\text{Cov}[\Upsilon_{\text{gg}}(R, R_0), \Upsilon_{\text{gg}}(R', R_0)]}{\Upsilon_{\text{gg}}(R, R_0) \Upsilon_{\text{gg}}(R', R_0)} + \frac{\sigma_\beta^2}{\beta^2}, \end{aligned} \quad (57)$$

where σ_β is the error in the measurement of β . This neglect of correlations is an approximation, justified in the case of our dataset by the fact that the sky area used for the galaxy clustering measurement is substantially different to the sub-sample used for galaxy-galaxy lensing (see Joachimi et al. (in prep.) for a detailed justification of this approximation), and that the projected lens clustering measurement (Υ_{gg}) is largely insensitive to redshift-space distortions (β) owing to the projection over the line-of-sight separations. We note that in our fiducial fitting approach, we determined the scale-independent statistic $\langle E_G \rangle$ through direct fits to Υ_{gm} and Υ_{gg} as discussed in Sect. 9, without requiring the covariance of $E_G(R)$.

5.4. Modification of noise term

We can replace the noise terms in Sects. 5.1 and 5.2 with a more accurate computation using the survey source and lens distributions. Neglecting the random lens term (which is not important on the small scales for which the noise term is significant), we

find that the variance associated with the γ_t estimator in Eq. 27 is (e.g. Miyatake et al. 2019),

$$\text{Var}[\gamma_t(\theta)] = \frac{\sum_{\text{ls}} w_l^2 w_s^2 \sigma_e^2}{\left(\sum_{\text{rs}} w_r w_s \right)^2}. \quad (58)$$

Likewise, the variance associated with the $\Delta\Sigma$ estimator in Eq. 30 is,

$$\text{Var}[\Delta\Sigma(R)] = \frac{\sum_{\text{ls}} w_l^2 w_s^2 w_{\text{ls}}^2 \sigma_e^2 (\Sigma_{\text{c,ls}})^2}{\left(\sum_{\text{rs}} w_r w_s w_{\text{rs}} \right)^2}. \quad (59)$$

We adopted these noise terms in our covariance model.

5.5. Modification for survey window

Eqs. 45 and 47 for the analytical covariance are modified by the survey window function. We can intuitively understand the need for this modification by considering that, whilst Fourier transforms assume periodic boundary conditions, the boundaries of the survey restrict the number of source-lens pairs on scales that are a significant fraction of the survey dimensions.

In Appendix B we derive how the covariance of a cross-correlation function $\xi(r)$ between two Gaussian fields is modified by the window function of the fields, $W_1(\mathbf{x})$ and $W_2(\mathbf{x})$ (see also, Beutler et al. 2017). We find,

$$\begin{aligned} \text{Cov}[\xi(r), \xi(s)] &\approx \frac{A_3(r, s)}{A_2(r) A_2(s)} \\ \frac{1}{2\pi} \int dk k [P_{11}(k) P_{22}(k) + P_{12}^2(k)] J_0(kr) J_0(ks), \end{aligned} \quad (60)$$

where P_{11} , P_{22} and P_{12} are the auto- and cross-power spectra of the fields and the pre-factors are given by,

$$\begin{aligned} A_2(r) &= \int_{\text{bin } r} d^3 \mathbf{r} \int d^2 \mathbf{x} W_1(\mathbf{x}) W_2(\mathbf{x} + \mathbf{r}) \\ A_3(r, s) &= \int_{\text{bin } r} d^3 \mathbf{r} \int_{\text{bin } s} d^3 \mathbf{s} \int d^2 \mathbf{x} A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{x}, \mathbf{s}), \end{aligned} \quad (61)$$

where the integrals over \mathbf{r} and \mathbf{s} are performed within the separation bin, and we have written $A_{12}(\mathbf{x}, \mathbf{r}) = W_1(\mathbf{x}) W_2(\mathbf{x} + \mathbf{r})$. We hence approximated the dependence of the covariance on the survey window by replacing the survey area in Eqs. 45 and 47 by the expression $A_2(r) A_2(s) / A_3(r, s)$.

We calculated the terms A_2 and A_3 using the mean and covariance of the pair count $R_s R_l(r)$ between random source and lens realisations (Landy & Szalay 1993), which have respective densities \bar{n}_s and \bar{n}_l . The mean pair count in a separation bin at scale r (between r_1 and r_2), containing bin area $A_{\text{bin}}(r) = \pi(r_2^2 - r_1^2)$, is

$$\langle R_s R_l(r) \rangle = \bar{n}_s \bar{n}_l A_{\text{bin}}(r) A_2(r), \quad (62)$$

which allows us to find $A_2(r)$, given that the other variables are known. The covariance of the pair count between separation bins r and s is,

$$\begin{aligned} \text{Cov}[R_s R_l(r), R_s R_l(s)] &= \\ \bar{n}_s \bar{n}_l A_{\text{bin}}(r) [A_2(r) \delta_{rs}^K + (\bar{n}_s + \bar{n}_l) A_{\text{bin}}(s) A_3(r, s)], \end{aligned} \quad (63)$$

which allows us to determine $A_3(r, s)$. For all the source-lens configurations and separation bins considered in this study, we find that the area correction factor differs from 1.0 by less than 10%.

5.6. Propagation of errors in multiplicative corrections

Galaxy-galaxy lensing measurements are subject to multiplicative correction factors arising from shape measurement calibration (see Sect. 6.1) and, in the case of $\Delta\Sigma$, owing to photometric redshift dilution (see Sect. 3.3). We propagated the uncertainties in these correction factors, which are correlated between different source and lens samples, into the analytical covariance of the measurements. Taking $\Delta\Sigma$ as an example and writing a general amplitude correction factor as α , the relation between the corrected and analytical statistics (denoted by the superscripts ‘corr’ and ‘ana’, respectively) is,

$$\Delta\Sigma_{ijk}^{\text{corr}} = \frac{\alpha_{ij} \Delta\Sigma_{ijk}^{\text{ana}}}{\langle\alpha_{ij}\rangle}, \quad (64)$$

which is normalised such that $\langle\Delta\Sigma_{ijk}^{\text{corr}}\rangle = \Delta\Sigma_{ijk}^{\text{ana}}$, where i denotes the lens sample, j the source sample and k the separation bin. We hence find,

$$\begin{aligned} \text{Cov}[\Delta\Sigma_{ijk}^{\text{corr}}, \Delta\Sigma_{lmn}^{\text{corr}}] &= \text{Cov}[\Delta\Sigma_{ijk}^{\text{ana}}, \Delta\Sigma_{lmn}^{\text{ana}}] (1 + C_{ij,lm}) \\ &\quad + \langle\Delta\Sigma_{ijk}^{\text{ana}}\rangle\langle\Delta\Sigma_{lmn}^{\text{ana}}\rangle C_{ij,lm}, \end{aligned} \quad (65)$$

where,

$$C_{ij,lm} = \frac{\langle\alpha_{ij}\alpha_{lm}\rangle - \langle\alpha_{ij}\rangle\langle\alpha_{lm}\rangle}{\langle\alpha_{ij}\rangle\langle\alpha_{lm}\rangle} = \frac{\text{Cov}[\alpha_{ij}, \alpha_{lm}]}{\langle\alpha_{ij}\rangle\langle\alpha_{lm}\rangle}. \quad (66)$$

We describe our specific implementation of these equations in the case of the KiDS dataset in Sect. 9.

6. Data

6.1. KiDS-1000

The Kilo-Degree Survey is a large optical wide-field imaging survey optimised for weak gravitational lensing analysis, performed with the OmegaCAM camera on the VLT Survey Telescope at the European Southern Observatory’s Paranal Observatory. The survey covers two regions of sky each containing several hundred square degrees, KiDS-North and KiDS-South, in four filters (u, g, r, i). The companion VISTA-VIKING survey has provided complementary imaging in near-infrared bands (Z, Y, J, H, K_s), resulting in a deep, wide, nine-band imaging dataset.

Our study is based on the fourth public data release of the project, KiDS-1000 (Kuijken et al. 2019), which comprises 1006 deg² of multi-band data, more than doubling the previously-available coverage. We used an early-science release of the KiDS-1000 shear catalogues, which was created using the exact pipeline version and PSF modelling strategy implemented in Hildebrandt et al. (2017) for the KiDS-450 release. We note that these catalogues have not undergone any rigorous assessment for the presence of cosmic shear systematics, but they are sufficient for the galaxy-galaxy lensing science presented in this paper, as this is less susceptible to systematic errors in the lensing catalogues. The raw pixel data was processed by the THELI and ASTRO_WISE pipelines (Erben et al. 2013; de Jong et al. 2015), and source ellipticities were measured using lensfit (Miller et al. 2013), assigning an optimal weight for each source, and calibrated by a large suite of image simulations (Kannawadi et al. 2019). Photometric redshifts z_B were determined from the nine-band imaging for each source using the Bayesian code BPZ (Benítez 2000), calibrated using spectroscopic sub-samples (Hildebrandt et al. 2020), and used to divide the sources into tomographic bins according to the value of z_B .

6.2. BOSS

The Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2013) is the largest existing galaxy redshift survey, which was performed using the Sloan Telescope between 2009 and 2014. BOSS mapped the distribution of 1.5 million Luminous Red Galaxies (LRGs) and quasars across $\sim 10\,000$ deg², inspiring a series of cosmological analyses including the most accurate existing measurements of baryon acoustic oscillations and redshift-space distortions in the galaxy clustering pattern (Alam et al. 2017a). The final (Data Release 12) large-scale structure catalogues are described by Reid et al. (2016); we used the combined LOWZ and CMASS LRG samples in our study.²

6.3. 2dFlenS

The 2-degree Field Lensing Survey (2dFlenS, Blake et al. 2016a) is a galaxy redshift survey performed at the Australian Astronomical Observatory in 2014-2015 using the 2-degree Field spectroscopic instrument, with the goal of extending spectroscopic-redshift coverage of gravitational lensing surveys in the southern sky, particularly the KiDS-South region. The 2dFlenS sample covers an area of 731 deg² and includes redshifts for 40 531 LRGs in the redshift range $z < 0.9$, selected by applying BOSS-inspired colour-magnitude cuts to the VST-ATLAS imaging data.³ The 2dFlenS dataset has already been utilised in conjunction with the KiDS-450 lensing catalogues to perform a previous implementation of the amplitude-ratio test (Amon et al. 2018), a combined cosmological analysis of cosmic shear tomography, galaxy-galaxy lensing and galaxy multipole power spectra (Joudaki et al. 2018) and to determine photometric redshift calibration by cross-correlation (Johnson et al. 2017; Hildebrandt et al. 2020). In our study we utilised the 2dFlenS LRG sample which overlapped with the KiDS-1000 pointings in the southern region. Fig. 1 illustrates the overlaps of the KiDS-1000 source catalogues in the north and south survey regions with the BOSS and 2dFlenS LRG catalogues.

7. Mocks

We used the MICECATv2.0 simulation (Fosalba et al. 2015b; Crocce et al. 2015; Fosalba et al. 2015a) to produce representative KiDS lensing source catalogues and LRG lens catalogues for testing the estimators, models and covariances described above. The Marenstrum Institut de Ciències de l’Espai (MICE) catalogues cover an octant of the sky ($0 < \text{RA} < 90^\circ$, $0 < \text{Dec} < 90^\circ$) for redshift range $z < 1.4$. We used boundaries at constant RA and Dec to divide this area into 10 sub-samples, each of area 516 deg². The fiducial set of cosmological parameters for the mock is $\Omega_m = 0.25$, $h = 0.7$, $\Omega_b = 0.044$, $\sigma_8 = 0.8$ and $n_s = 0.95$.

7.1. Mock source catalogue

We constructed the representative mock source catalogue by applying the following steps (see van den Busch et al. (in prep.) for a full description of the MICE KiDS source mocks). The MICE catalogue is non-uniform across the octant: the region $\text{Dec} < 30^\circ$ AND $[(\text{RA} < 30^\circ) \text{ OR } (\text{RA} > 60^\circ)]$ has a shallower

² The BOSS large-scale structure samples are available for download at the link <https://data.sdss.org/sas/dr12/boos/lss/>.

³ The 2dFlenS dataset is publicly available at the link <http://2df lens.swin.edu.au>.

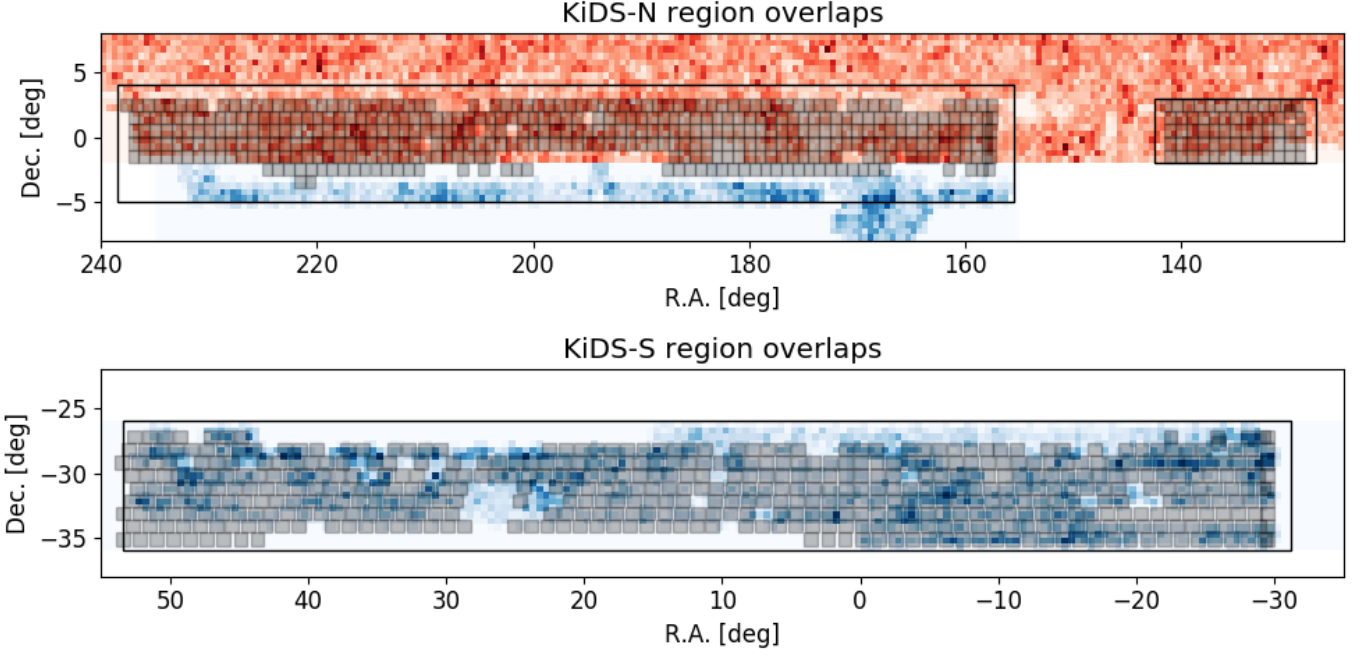


Fig. 1. KiDS source and LRG lens catalogues within the KiDS-N region (top panel) and KiDS-S region (bottom panel). The grey squares represent the KiDS-1000 pointings, and the fluctuating background is the gridded number density of 2dFLenS (blue) and BOSS (red) galaxies. The open rectangles outline the footprint of the full KiDS survey, indicating that the LRG overlap will continue to increase as the survey is completed.

redshift distribution than the remainder. Firstly, we homogenised the catalogue with the cut `des_asahi_full_i_true < 24`, such that we could construct mocks using the complete octant. The MICE catalogue shears (γ_1, γ_2) are defined by the position angle relative to the declination axis. Given the MICE system for mapping 3D positions to (RA, Dec) co-ordinates, the KiDS conventions can be recovered by the following transformations: $RA \rightarrow 90^\circ - RA$, $\gamma_1 \rightarrow -\gamma_1$ (γ_2 is effectively negated twice and therefore unchanged).

We constructed a KiDS-like photometric realisation based on the galaxy sizes and shapes, median KiDS seeing and limiting magnitudes, including photometric noise (see van den Busch et al. in prep.). We ran BPZ photometric redshift estimation (Benítez 2000) on the mock source magnitudes and sizes, assigning z_B values for each object. We used a KDTree algorithm to assign weights to the mock sources on the basis of a nearest-neighbour match to the data catalogue in magnitude space, and randomly sub-sampled the catalogue to match the KV450 effective source density. We produced noisy shear components (e_1, e_2) as $e = (\gamma + n)/(1 + n\gamma^*)$ (Seitz & Schneider 1997) where $\gamma = \gamma_1 + \gamma_2 i$, $e = e_1 + e_2 i$ and $n = n_1 + n_2 i$, where n_1 and n_2 are drawn from Gaussian distributions with standard deviation $\sigma_e = 0.288$ (Hildebrandt et al. 2020). The redshift distribution estimates of the KiDS data and MICE mock source tomographic samples are displayed in the left panel of Fig. 2, illustrating the reasonable match between the two catalogues.

7.2. Mock lens catalogue

We constructed the representative mock LRG lens catalogue from the MICE simulation as follows. We used the galaxy magnitudes `sdss_g_true`, `sdss_r_true`, `sdss_i_true` and first applied the MICE evolution correction to these magnitudes as a function of redshift, $m \rightarrow m - 0.8 * [\arctan(1.5 * z) - 0.1489]$ (Crocce et al. 2015). We then constructed the LRG lens cata-

logues using the BOSS LOWZ and CMASS colour cuts in terms of the variables,

$$\begin{aligned} c_{\parallel} &= 0.7(g - r) + 1.2(r - i - 0.18), \\ c_{\perp} &= (r - i) - (g - r)/4 - 0.18, \\ d_{\perp} &= (r - i) - (g - r)/8. \end{aligned} \quad (67)$$

Applying the original BOSS colour-magnitude selection cuts (Eisenstein et al. 2011) to the MICE mock did not reproduce the BOSS redshift distribution (which is unsurprising, since this mock has not been tuned to do so; the BOSS data is also selected from noisy observed magnitudes). Our approach to resolve this issue, following Crocce et al. (2015), was to vary the colour and magnitude selection cuts to minimise the deviation between the mock and data redshift distributions. We applied the following LOWZ selection cuts (where we indicate our changed values in bold font, and the previous values immediately following in square brackets):

$$\begin{aligned} 16.0 < r < \mathbf{20.0} [19.6], \\ r < \mathbf{13.35} [13.5] + c_{\parallel}/0.3, \\ |c_{\perp}| < 0.2. \end{aligned} \quad (68)$$

We applied the following CMASS selection cuts:

$$\begin{aligned} 17.5 < i < \mathbf{20.06} [19.9], \\ r - i < 2, \\ d_{\perp} > 0.55, \\ i < \mathbf{19.98} [19.86] + 1.6(d_{\perp} - 0.8). \end{aligned} \quad (69)$$

The resulting redshift distributions of the MICE lens mock (original, and after adjustment of the colour selection cuts) and the BOSS data are shown in the right panel of Fig. 2, illustrating that our modified selection produced a much-improved representation of the BOSS dataset. The clustering amplitude of the

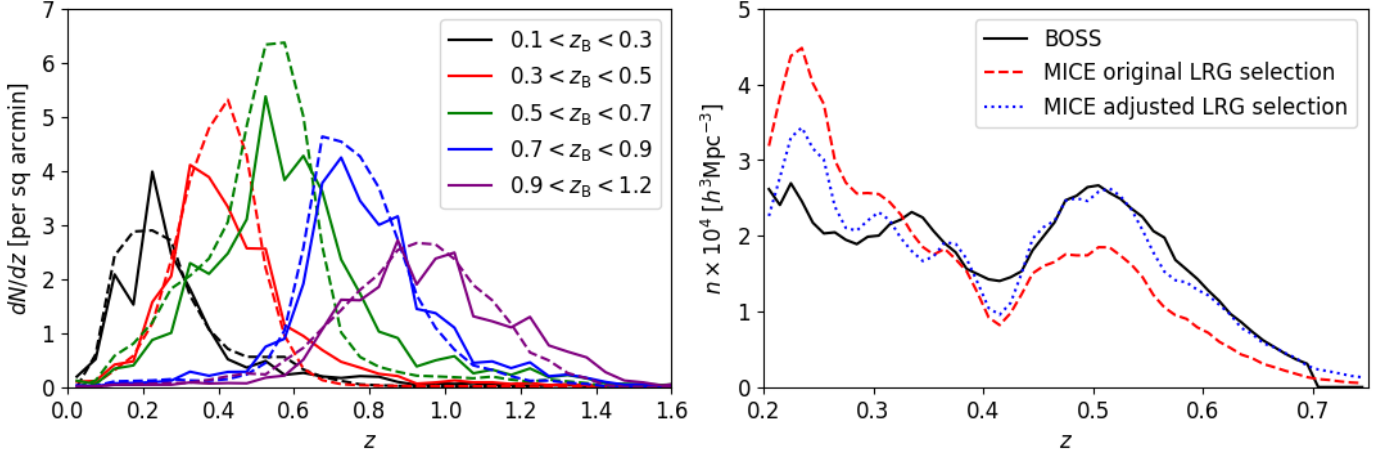


Fig. 2. *Left panel:* Redshift distribution estimates of the KiDS source catalogue (solid lines) and MICE mock source catalogue (dashed lines) in the five source tomographic bins split by photometric redshift. *Right panel:* Number density as a function of redshift of the BOSS LRG dataset (black solid line), the MICE mock lens catalogue with the original BOSS colour selection (red dashed line), and the MICE mock lens catalogue with the adjusted BOSS colour selection (blue dotted line).

MICE LRG mock catalogues was consistent with a galaxy bias factor $b \approx 2$, although did not precisely match the clustering of the BOSS dataset, since it was not tuned to do so. However, these representative catalogues nonetheless allowed us to test our analysis procedures.

8. Simulation tests

In this section we analyse the representative source and lens catalogues constructed from the MICE mocks described in Sect. 7. Our specific goals are to:

- Test that the non-linear galaxy bias model specified in Sect. 2.4 is adequate for modelling the galaxy-galaxy lensing and clustering statistics across the relevant scales.
- Test that the approaches to the photo- z dilution correction of $\Delta\Sigma$ described in Sect. 3.3 recovered results consistent with those obtained using source spectroscopic redshifts.
- Use the multiple mock realisations and jack-knife techniques to test that the covariance of the estimated statistics is consistent with the analytical Gaussian covariance specified in Sect. 5.
- Test that the E_G test statistics constructed from the mock as described in Sect. 4.3 are consistent with the theoretical expectation, and determine the degree to which this result depends on the choice of the small-scale cut-off parameter R_0 (see Eq. 40).
- Test that, given our galaxy bias model, the galaxy-galaxy lensing and clustering statistics may be jointly described by a normalisation parameter σ_8 that is consistent with the mock fiducial cosmology, and use this test to assess the relative precision of angular and projected estimators.

Consistent with our subsequent data analysis, we divided the source catalogues into five different tomographic samples by the value of the BPZ photometric redshift, with divisions $z_B = [0.1, 0.3, 0.5, 0.7, 0.9, 1.2]$ (following Hildebrandt et al. 2020). We divided the lens catalogue into five slices of spectroscopic redshift z_1 of width $\Delta z_1 = 0.1$ in the range $0.2 < z_1 < 0.7$. This narrow spectroscopic slicing minimises systematic effects due to redshift evolution across the lens slice (Leauthaud et al. 2017; Singh et al. 2019).

8.1. Measurements

We measured the following three statistics, which form the fundamental set of measurements from which the associated statistics are derived. Firstly, we measured the average tangential shear $\gamma_t(\theta)$ between all tomographic pairs of source and lens samples, in 15 logarithmically-spaced angular bins in the range $0.005^\circ < \theta < 5^\circ$, using the estimator of Eq. 27. This measurement is displayed in Fig. 3 as the mean of the 10 individual mock realisations (which each have area 516 deg^2), for each of the five source samples against the five lens redshift slices.

Secondly, we measured the projected mass density $\Delta\Sigma(R)$ between all tomographic pairs of source and lens samples, in 15 logarithmically-spaced projected separation bins in the range $0.1 < R < 100 h^{-1} \text{ Mpc}$. The mock mean measurement is displayed in Fig. 4 in units of $h M_\odot \text{ pc}^{-2}$, for each of the five source samples against the five lens redshift slices. When performing a $\Delta\Sigma$ measurement between source and lens samples we only included individual source-lens galaxy pairs with $z_B > z_1$, for which the source photometric redshift lies behind the lens spectroscopic redshift (adopting an alternative cut $z_B > z_1 + 0.1$ did not change the results significantly). We applied the photo- z dilution correction f_{bias} computed using Eq. 35 based on the point photo- z values, and we study the efficacy of this correction in Sect. 8.2 below.

Thirdly, we measured the projected clustering $w_p(R)$ of the lens samples in the same projected separation bins as above, using the estimator of Eq. 39 with $\Pi_{\text{max}} = 100 h^{-1} \text{ Mpc}$. The mock mean measurement of $w_p(R)$ is displayed as the third row of Fig. 5, for each of the five redshift slices. Each of these measurements will be compared with cosmological model predictions as described in the subsections below.

We computed the covariance matrix for each statistic using the analytical Gaussian covariance specified in Sect. 5, where we initially used a fiducial lens linear bias factor $b_L = 1.8$, and iterated this value following a preliminary fit to the projected lens clustering. Fig. 6 compares three different determinations of the error in $\Delta\Sigma(R)$ for each individual 516 deg^2 realisation of the MICE mocks: using the analytical covariance, using a jack-knife analysis, and derived from the standard deviation of the 10 realisations. For the jack-knife analysis, we divided the sample into 7×7 angular regions using constant boundaries in RA and

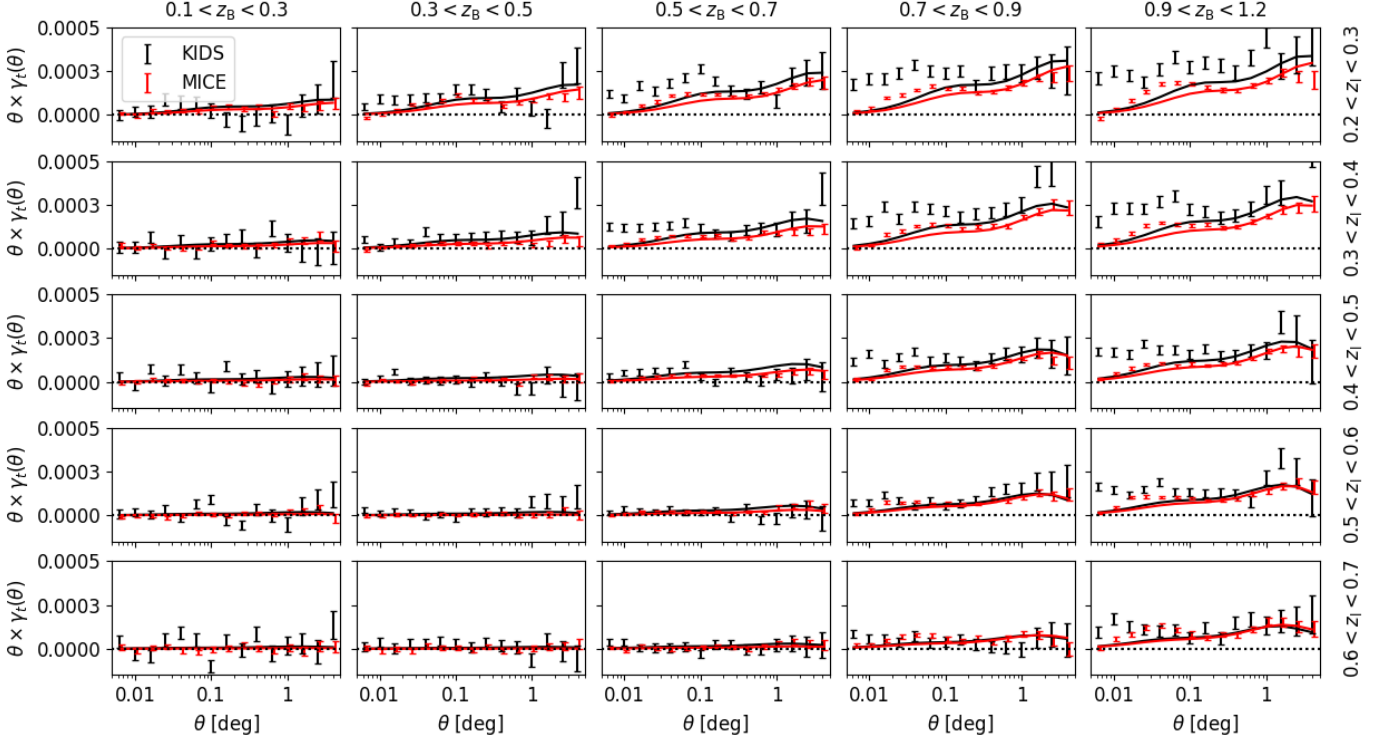


Fig. 3. Measurements of the average tangential shear, $\gamma_t(\theta)$, for the KiDS-LRG dataset (black points) and representative MICE mocks (red points), between all pairs of lens spectroscopic redshift slices (rows) and source tomographic samples split by photometric redshift (columns). The errors are derived from the diagonal elements of the full analytical covariance matrix (where we note that measurements are correlated across scales and samples). The overplotted model is not a fit to this dataset, but rather a prediction based on the galaxy bias parameter fits to the Υ_{gg} statistic of each lens redshift slice, which is inaccurate on small scales. The y -values are scaled by a factor of θ (in degrees) for clarity.

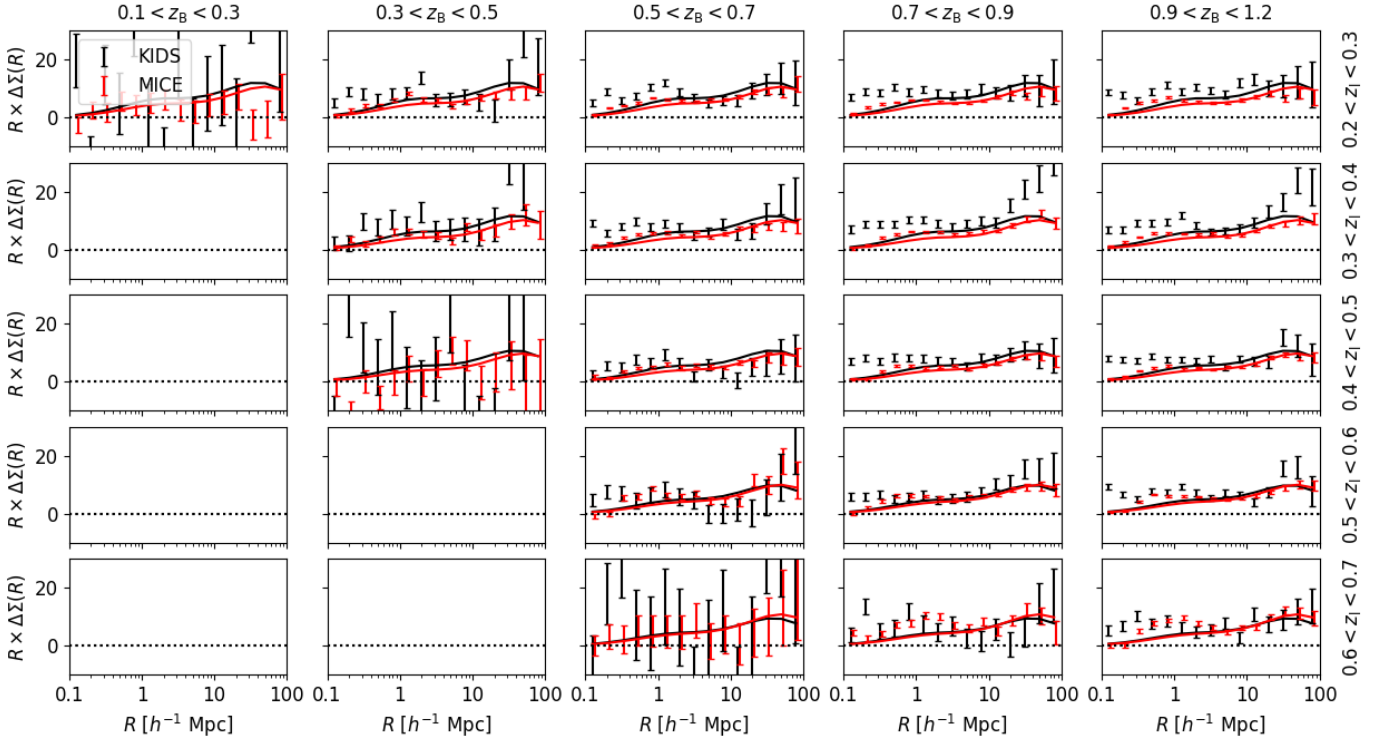


Fig. 4. Measurements of the projected mass density, $\Delta\Sigma(R)$, for the KiDS-LRG dataset (black points) and representative MICE mocks (red points), displayed in the same style as Fig. 3. No measurements are possible for the lower left-hand set of panels, owing to the adopted cut in source-lens pairs, $z_B > z_1$. The units of $\Delta\Sigma$ are $h \text{ M}_\odot \text{ pc}^{-2}$.

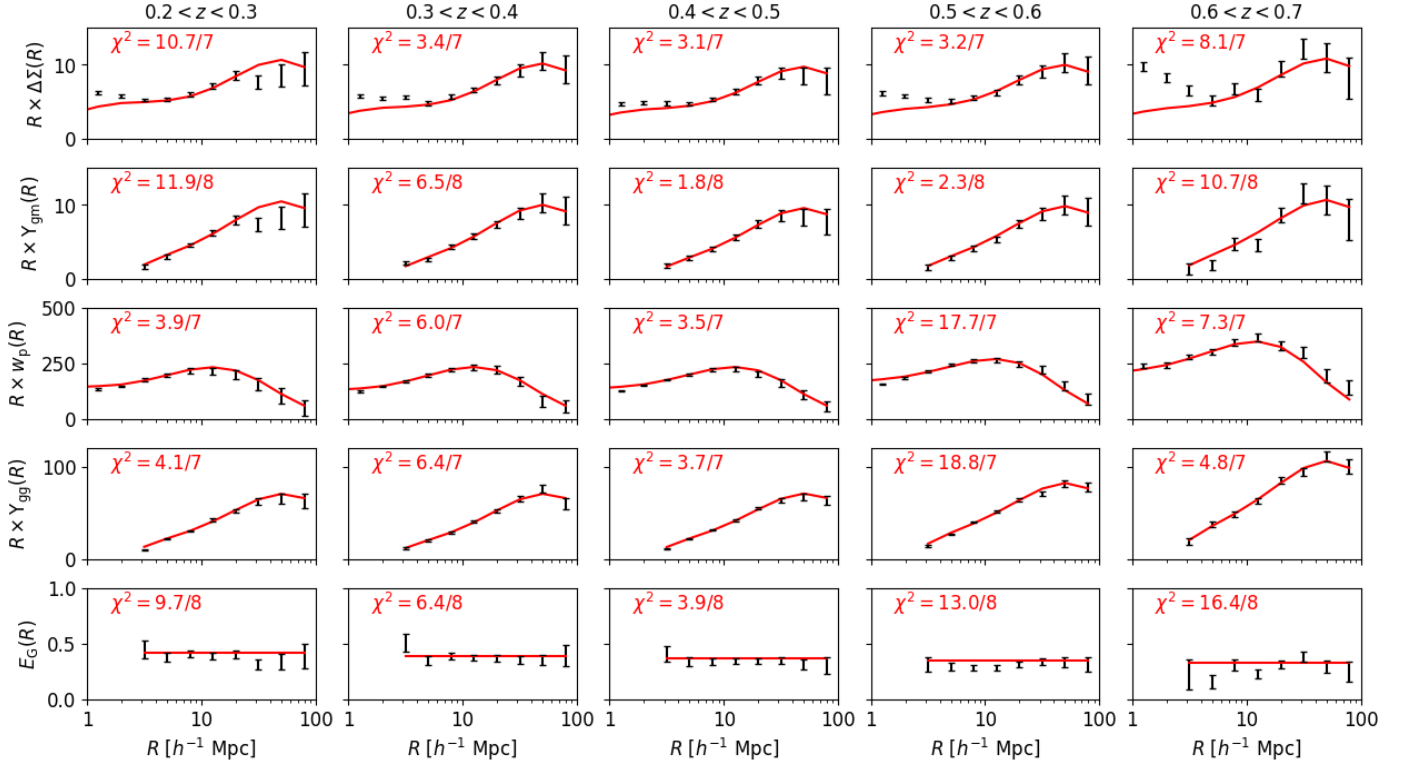


Fig. 5. Measurements of a series of galaxy-galaxy lensing and clustering statistics (rows) for each different lens redshift slice (columns) of the MICE mocks. The first row displays the projected mass density $\Delta\Sigma(R)$, where measurements corresponding to the different source tomographic samples have been optimally combined following the procedure described in Appendix C. The second row shows the corresponding measurements of $\Upsilon_{\text{gm}}(R)$, derived from $\Delta\Sigma(R)$ using Eq. 40 and assuming $R_0 = 2 h^{-1}$ Mpc. The third and fourth rows display the galaxy clustering statistics $w_p(R)$ and $\Upsilon_{\text{gg}}(R)$ for each lens redshift slice, where $w_p(R)$ is the projected correlation function measurement, and $\Upsilon_{\text{gg}}(R)$ is the associated quantity suppressing contributions from small scales, derived using Eq. 41 and again assuming $R_0 = 2 h^{-1}$ Mpc. The fifth row shows the combined-probe statistic $E_G(R)$, derived from the $\Upsilon_{\text{gm}}(R)$ and $\Upsilon_{\text{gg}}(R)$ measurements using Eq. 44. Errors are displayed using the diagonal elements of the analytical covariance matrix, propagating errors where appropriate. The overplotted models are determined using the galaxy bias factors fitted to the Υ_{gg} measurements for each lens redshift slice, and χ^2 statistics between the mock mean data and model are displayed in each panel.

Dec, such that each region contained the same angular area 10.5 deg^2 . In Fig. 6 we display the comparison as a ratio between the jack-knife or realisations error, and the analytical error.

We find that in the range $R > 1 h^{-1}$ Mpc, where the model provides a reasonable description of the measurements, the average (fractional) absolute difference between the analytical and jack-knife errors is 15%, and between the analytical and realisation scatter is 21% (which is the expected level of difference given the error in the variance for 10 realisations). Small differences between these error estimates may arise due to the Gaussian approximation in the analytical covariance, the exact details of the survey modelling, or the scale of the jack-knife regions.

Fig. 7 displays the full analytical covariance matrix of $\Delta\Sigma(R)$ – spanning five lens redshift slices, five source tomographic samples and 15 bins of scale – as a correlation matrix with 375×375 entries. We note that there are significant off-diagonal correlations between measurements utilising the same lens or source sample, and between different scales. The covariance matrix is reduced in dimension if source-lens sample pairs with $z_B < z_L$ are excluded, as illustrated by the missing panels in Figs. 4.

We combined the correlated $\Delta\Sigma$ measurements for each individual lens redshift slice, averaging over the five different source tomographic samples, using the procedure described in Appendix C. The resulting combined $\Delta\Sigma$ measurement for each lens redshift sample (again corresponding to a mock mean) is shown as the first row in Fig. 5.

We then used the $\Delta\Sigma(R)$ and $w_p(R)$ measurements to infer the Upsilon statistics, $\Upsilon_{\text{gm}}(R, R_0)$ and $\Upsilon_{\text{gg}}(R, R_0)$, using Eqs. 40 and 41 respectively, adopting a fiducial value $R_0 = 2 h^{-1}$ Mpc (we consider the effect of varying this choice below). These measurements are shown in the second and fourth rows of Fig. 5. We determined the covariance of $\Upsilon_{\text{gm}}(R, R_0)$ and $\Upsilon_{\text{gg}}(R, R_0)$ using error propagation following Eqs. 55 and 56, respectively.

Finally, we determined the $E_G(R)$ statistic for each lens redshift slice using Eq. 44 where, for the purposes of these tests focussed on galaxy-galaxy lensing, we assumed a fixed input value for the redshift-space distortion parameter $\beta = f(z)/b_L(z)$, where we evaluated $f(z) = \Omega_m(z)^{0.55}$ using the fiducial cosmology of the MICE simulation – we note that the exponent 0.55 is an excellent approximation to the solution of the differential growth equation in Λ CDM cosmologies (Linder 2005) – and $b_L(z)$ is the best-fitting linear bias parameter to the Υ_{gg} measurements for each lens redshift slice z . Hence, systematic errors associated with redshift-space distortions lie beyond the scope of this study, and in our subsequent data analysis we will infer the required β values from existing literature. We propagated errors in E_G using Eq. 57 (and assuming no error in β in the case of the mocks). Our E_G measurements are shown as the fifth row in Fig. 5.

We generated fiducial cosmological models for these statistics using a non-linear matter power spectrum $P(k, z)$ corresponding to the fiducial cosmological parameters of the MICE simulation listed in Sect. 7. We determined the best-fitting lin-

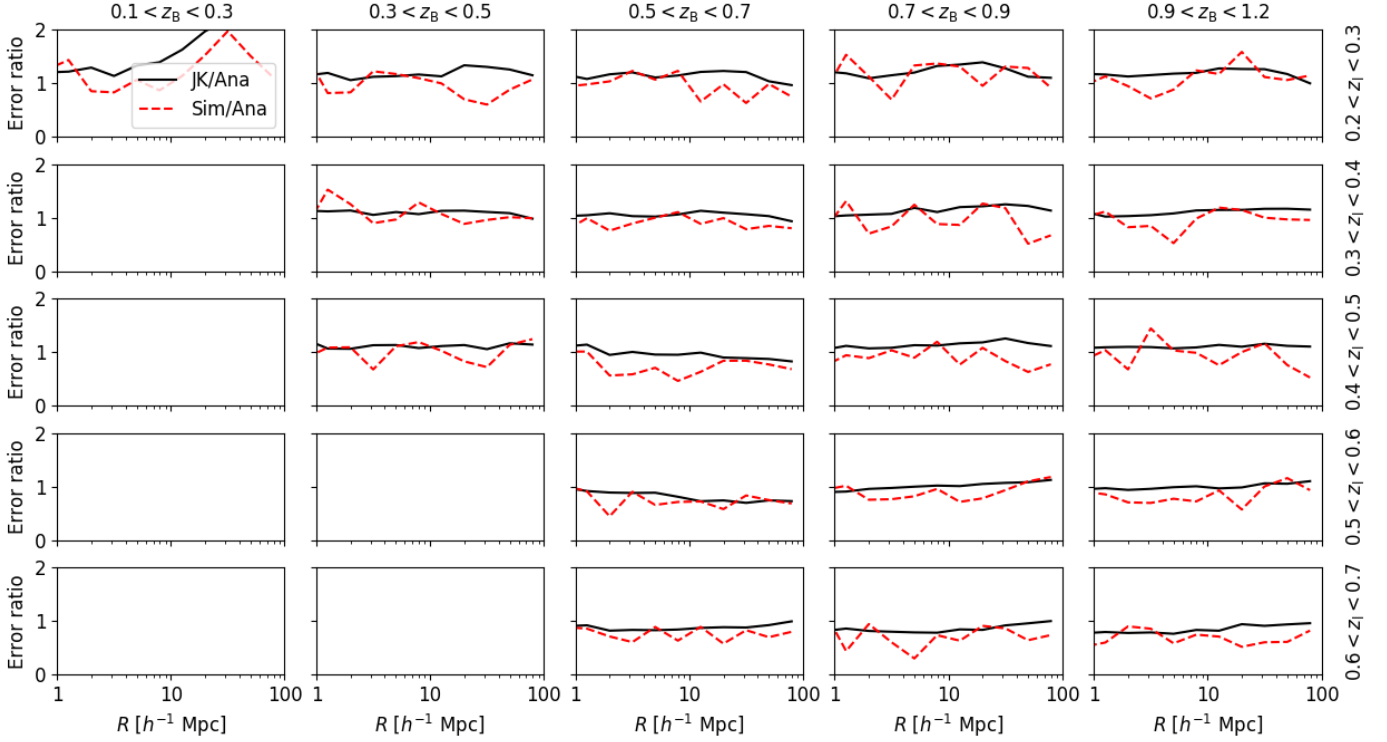


Fig. 6. Comparison of different estimates of the error in $\Delta\Sigma(R)$ for individual 516 deg^2 realisations of the MICE mocks, between all pairs of lens spectroscopic redshift slices (rows) and source tomographic samples (columns). The black solid line shows the ratio between the error determined by a jack-knife analysis of the data and the error derived from the diagonal elements of the analytical covariance matrix, and the red dashed line is the ratio between the standard deviation across 10 realisations and the analytical error. No measurements are possible for the lower left-hand set of panels, owing to the adopted cut in source-lens pairs, $z_B > z_1$.

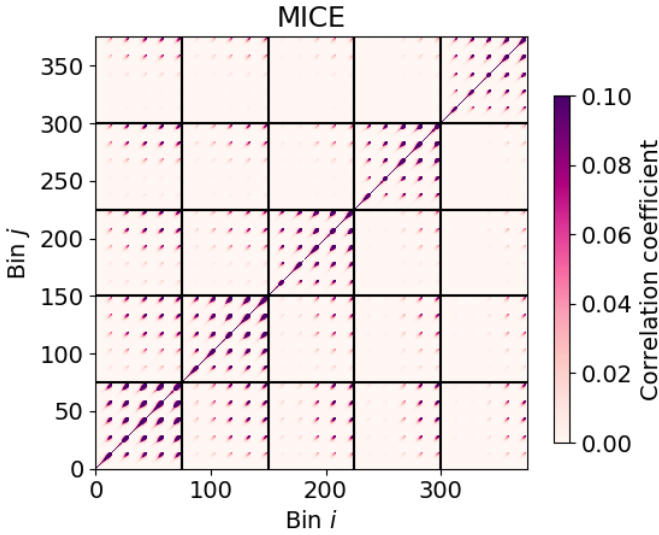


Fig. 7. Full analytical covariance matrix C_{ij} of the projected mass density $\Delta\Sigma(R)$ for the MICE mocks, spanning five lens redshift slices, five source tomographic samples and 15 bins of scale – i.e. 375×375 entries – ordered where each subsequent bin changes separation, then source sample, then lens sample, such that the vertical and horizontal solid lines demarcate different lens redshift slices. We display the results as a correlation matrix $r = C_{ij} / \sqrt{C_{ii} C_{jj}}$, and note that there are significant off-diagonal correlations between measurements utilising the same lens or source sample, and between different scales. We note that the colour bar is saturated at $r = 0.1$ to reveal low-amplitude cross-correlation more clearly.

ear and non-linear galaxy bias parameters (b_L, b_{NL}) by fitting to the Υ_{gg} measurements for each lens redshift slice for scales $R > 5 h^{-1} \text{ Mpc}$, and applied these same bias parameters to the galaxy-galaxy lensing models. The models plotted in Figs. 3, 4 and 5 do not otherwise contain any free parameters. In Fig. 5 we display corresponding χ^2 statistics between the models and mock mean data, demonstrating a satisfactory goodness-of-fit in general. We evaluated the χ^2 statistics for $R > 5 h^{-1} \text{ Mpc}$ for $\Delta\Sigma$, w_p and Υ_{gg} , and using all scales for Υ_{gm} and E_G . We conclude that our lensing and clustering measurements from the MICE mocks generally agree with the underlying ΛCDM cosmology, which we further explore via cosmological parameter fitting in Sect. 8.3.

8.2. Photo- z dilution correction

Within our mock analysis we considered three different implementations of the photo- z dilution correction necessary for the $\Delta\Sigma(R)$ measurements, as described in Sect. 3.3. Firstly, we used the source spectroscopic redshift values (which are available given that this is a simulation) to produce a baseline $\Delta\Sigma$ measurement free of photo- z dilution. Secondly, for our fiducial analysis choice, we used the source photometric redshift point values in the estimator of Eq. 31, adopting a source-lens pair cut $z_B > z_1$ and correcting for the photo- z dilution using the f_{bias} factor of Eq. 35. We also considered the same case, excluding the f_{bias} correction factor. Thirdly, we used the redshift probability distributions for each source tomographic sample to determine Σ_c^{-1} relative to each lens redshift using Eq. 11, and then estimated

$\Delta\Sigma$ using Eq. 37. We refer to this as the $P(z)$ distribution-based method.

The results of these $\Delta\Sigma$ analyses are compared in Fig. 8 for each lens redshift slice, where measurements corresponding to the different source tomographic samples have been optimally combined. We find that, other than in the case where the f_{bias} correction is excluded, both the point-based and distribution-based photo- z dilution corrections produce $\Delta\Sigma$ measurements which are statistically consistent with the baseline measurements using the source spectroscopic redshifts. We further verify in Sect. 8.3 that these analysis choices do not create significant differences in cosmological parameter fits.

8.3. Recovery of cosmological parameters

Finally, we verified that our analysis methodology recovered the fiducial cosmological parameters of the MICE simulation within an acceptable statistical accuracy. In this study we focus only on the amplitudes of the clustering and lensing statistics, keeping all other cosmological parameters fixed. In particular we test the recovery of the E_G statistics, and the recovery of the σ_8 normalisation, marginalising over galaxy bias parameters.

First, we determined a scale-independent E_G value (which we denote $\langle E_G \rangle$) for each lens redshift slice from the MICE mock mean statistics displayed in Fig. 5. We considered two approaches to this determination. In one approach, we fitted a constant value to the $E_G(R)$ measurements shown in the fifth row of Fig. 5 using the corresponding analytical covariance matrix, that is, varying a vector of five parameters,

$$\mathbf{p} = [E_G(z_1), E_G(z_2), E_G(z_3), E_G(z_4), E_G(z_5)]. \quad (70)$$

This approach has the disadvantage that it is based on the ratio of two noisy quantities $\Upsilon_{\text{gm}}/\Upsilon_{\text{gg}}$, which may result in a biased or non-Gaussian result.

Our second approach avoided this issue by including $\langle E_G \rangle$ as an additional parameter in a joint fit to the Υ_{gm} and Υ_{gg} statistics for each lens redshift slice, where $\langle E_G \rangle$ changed the amplitude of Υ_{gm} relative to Υ_{gg} . Specifically, we fitted the model,

$$\begin{aligned} \Upsilon_{\text{gm}}(R) &= A_E b_L \Upsilon_{\text{gm}}(R, \sigma_8 = 0.8, b_L, b_{\text{NL}}) \\ \Upsilon_{\text{gg}}(R) &= \Upsilon_{\text{gg}}(R, \sigma_8 = 0.8, b_L, b_{\text{NL}}), \end{aligned} \quad (71)$$

in terms of an amplitude parameter A_E and galaxy bias parameters b_L and b_{NL} , and then determined $\langle E_G \rangle = A_E b_L E_{G,\text{fid}}$ for each lens redshift slice, where $E_{G,\text{fid}}(z) = \Omega_m/f(z)$ in terms of the fiducial matter density parameter of the MICE mocks, Ω_m , and the theoretical growth rate of structure based on this matter density, $f(z)$. Hence we vary a vector of 15 parameters,

$$\mathbf{p} = [A_{E,1}, b_{L,1}, b_{\text{NL},1}, \dots, A_{E,5}, b_{L,5}, b_{\text{NL},5}]. \quad (72)$$

We note that the b_L factor in Eq. 71 for Υ_{gm} arises as a consequence of our treatment of β as a fixed input parameter as described in Sect. 8.1, and ensures that A_E is constrained only by the relative ratio $\Upsilon_{\text{gm}}/\Upsilon_{\text{gg}}$, and not the absolute amplitude of these functions.

Fig. 9 displays the different determinations of $\langle E_G \rangle$ in each lens redshift slice. The left panel compares measurements using the four different treatments of photo- z dilution shown in Fig. 8, confirming that these methods produce consistent E_G determinations (other than the case in which f_{bias} is excluded; our fiducial choice is the direct photo- z pair counts with $z_B > z_1$). The middle panel compares $\langle E_G \rangle$ fits varying the small-scale parameter R_0 (where our fiducial choice is $R_0 = 2.0 h^{-1}$ Mpc, and we also considered choices corresponding to the adjacent separation bins 1.2

and $3.1 h^{-1}$ Mpc). The right panel alters the method used to determine $\langle E_G \rangle$, comparing the default choice using the non-linear bias model, a linear model where we fix $b_{\text{NL}} = 0$, and a direct fit to the scale-dependent $E_G(R)$ values. Reassuringly, all these methods yielded very similar results.

We compared these determinations to the model prediction $E_G(z) = \Omega_m/f(z)$ shown in Fig. 9. Other than for the case where the f_{bias} correction is excluded, both the point-based and distribution-based photo- z dilution corrections produce determinations of $\langle E_G \rangle$ which recover the fiducial value. This conclusion holds independently of the chosen value of R_0 , although higher R_0 values produce slightly increased error ranges. The different modelling approaches also produce consistent results.

Next, we utilised our mock dataset to perform a fit of the cosmological parameter σ_8 to the joint lensing and clustering statistics, marginalising over different bias parameters (b_L, b_{NL}) for each redshift slice such that we vary a vector of 11 parameters,

$$\mathbf{p} = [\sigma_8, b_{L,1}, b_{\text{NL},1}, \dots, b_{L,5}, b_{\text{NL},5}]. \quad (73)$$

We fixed the remaining cosmological parameters, and performed our parameter fit using a Markov Chain Monte Carlo method implemented using the *emcee* package (Foreman-Mackey et al. 2013). We used wide, uniform priors for each fitted parameter.

As above, we adopted for our fiducial analysis the point photo- z dilution correction using f_{bias} , and we performed fits to the $\Upsilon_{\text{gm}}(R)$ and $\Upsilon_{\text{gg}}(R)$ statistics with $R_0 = 2 h^{-1}$ Mpc, considering the same analysis variations as above. Following the scale cuts mentioned above, the data vector contains eight scales for $\Upsilon_{\text{gm}}(R)$ and seven scales for $\Upsilon_{\text{gg}}(R)$ for each of the five lens redshift slices, comprising a total of 75 data points.

For this fiducial case, we obtained a measurement $\sigma_8 = 0.779 \pm 0.019$, consistent with the MICE simulation cosmology $\sigma_8 = 0.8$. The χ^2 statistic of the best-fitting model is 69.9 for 64 degrees of freedom (d.o.f.), that is, 75 data points minus the 11 fitted parameters. Fig. 10 displays the dependence of the σ_8 measurements on the analysis choices. All methodologies using the non-linear bias model recovered the fiducial σ_8 value, with the exception of excluding the f_{bias} correction. Adopting a linear bias model instead produced a significantly poorer recovery.

We also considered fitting to different pairs of lensing-clustering statistics: $\Delta\Sigma(R)$ and $w_p(R)$ for $R > R_{\text{min}} = 5 h^{-1}$ Mpc, compared to $\gamma_t(\theta)$ and $w_p(R)$, where we applied a minimum-scale cut in θ which matches R_{min} in each lens redshift slice. These alternative statistics also successfully recovered the fiducial value of σ_8 , with errors of 0.018 (for $\Delta\Sigma$) and 0.022 (for γ_t). According to this analysis, the projected statistics produced a $\sim 20\%$ more accurate σ_8 value than the angular statistics, in agreement with the results of Shirasaki & Takada (2018).

We conclude this section by noting that the application of our analysis pipeline to the MICE lens and source mocks successfully recovered the fiducial E_G and σ_8 parameters of the simulation, and is robust against differences in photo- z dilution correction, choice of the small-scale parameter R_0 , and choice of statistic included in the analysis [$\gamma_t(\theta)$, $\Delta\Sigma(R)$ or $\Upsilon_{\text{gm}}(R)$].

9. Results

9.1. Measurements

We now summarise the galaxy-galaxy lensing and clustering measurements we generated from the KiDS-1000 and overlapping LRG datasets. We cut these catalogues to produce overlapping subsets for our galaxy-galaxy lensing analysis, by only

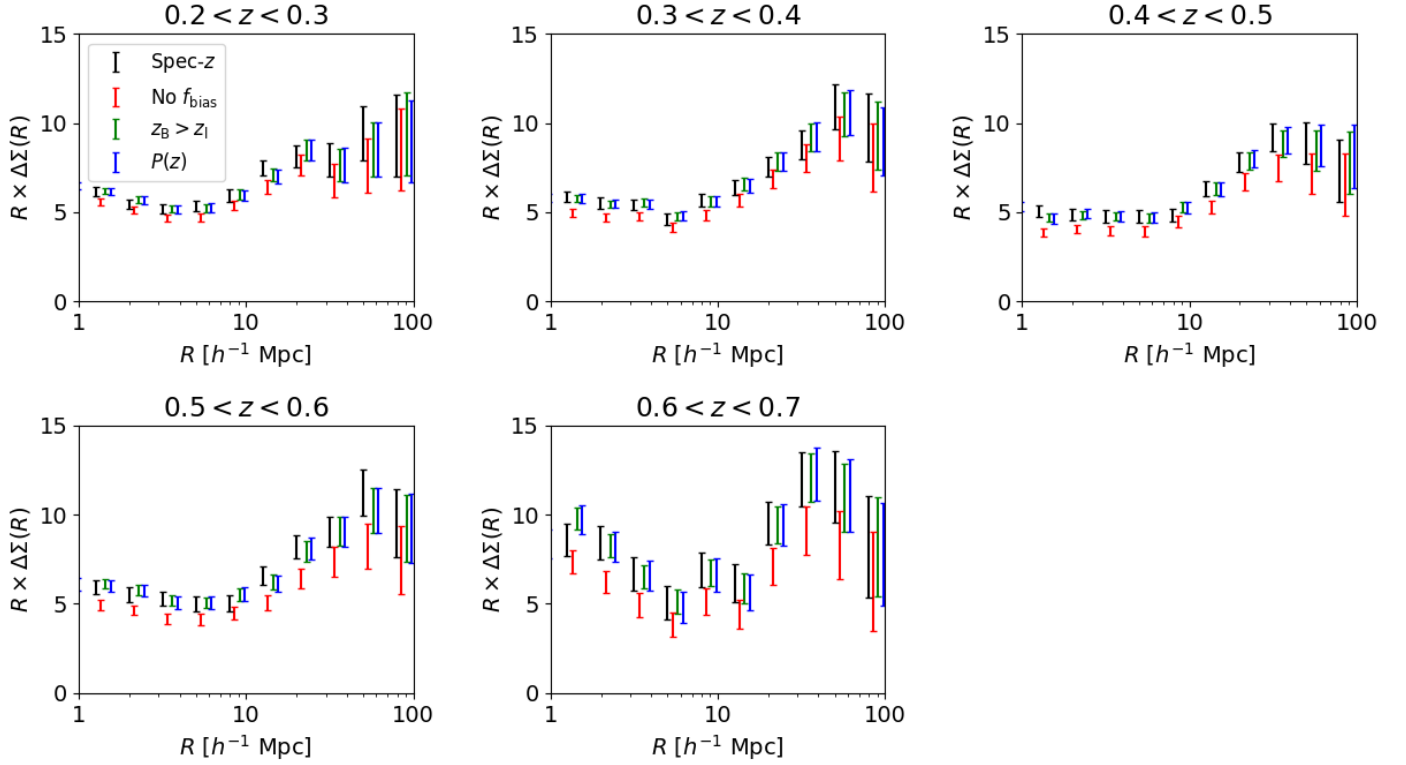


Fig. 8. Measurements of the projected mass density $\Delta\Sigma$ for each lens redshift slice of the MICE mocks, where measurements corresponding to the different source tomographic samples have been optimally combined following the procedure described in Appendix C. Results are shown for four cases: using the spectroscopic redshifts of the sources (black points, fiducial measurement with perfect source redshifts), using the photometric redshifts of the sources but without a correction for the photo- z dilution factor f_{bias} (red points, photo- z dilution remains uncorrected), including the dilution correction (green points), and using the redshift probability distribution for each source tomographic slice (blue points). We only include individual source-lens pairs with $z_B > z_l$ in the measurement. We find that, other than for the case where the f_{bias} correction is excluded, both the point-based and distribution-based photo- z dilution corrections produce $\Delta\Sigma$ measurements which are consistent with those obtained using the source spectroscopic redshifts.

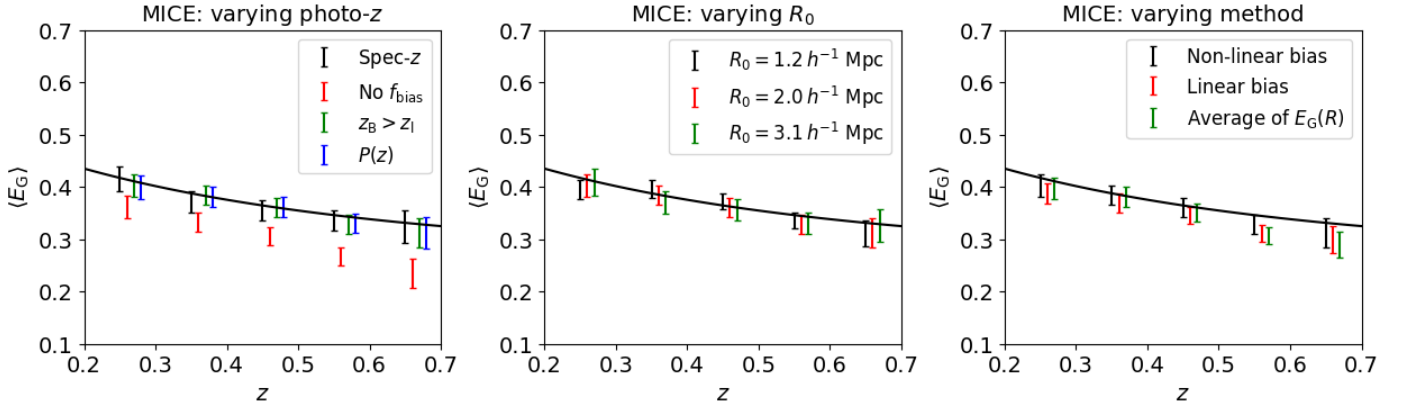


Fig. 9. Comparison of the scale-independent values $\langle E_G \rangle$ determined for each lens redshift slice of the MICE mocks, varying the galaxy-galaxy lensing analysis assumptions and methodology. Our fiducial analysis adopted a point-based photo- z correction with $z_B > z_l$, $R_0 = 2.0 h^{-1} \text{ Mpc}$ and a model fitted to $(Y_{\text{gm}}, Y_{\text{gg}})$ including non-linear galaxy bias. The left panel compares determinations of $\langle E_G \rangle$ varying the photo- z dilution correction method studying the same four cases described in Fig. 8, the middle panel varies the small-scale parameter R_0 , and the right panel alters the fitting method to only include linear galaxy bias, and to use a direct fit to the $E_G(R)$ values. The model line in each case is the prediction $E_G(z) = \Omega_m / f(z)$, where Ω_m is the fiducial matter density parameter for the MICE mocks.

retaining sources and lenses within the set of KiDS pointings which contain BOSS or 2dFLenS galaxies. The resulting KiDS-N sample comprised 15 150 250 KiDS shapes and 47 332 BOSS lenses within 474 KiDS pointings with total unmasked area 366.0 deg^2 , and the KiDS-S sample consisted of 16 994 252 KiDS shapes and 18 903 2dFLenS lenses within 478 KiDS point-

ings with total unmasked area 382.1 deg^2 . We also utilised BOSS and 2dFLenS random catalogues in our analysis, with the same selection cuts and size 50 times bigger than the datasets, subsampled from the master random catalogues provided by Reid et al. (2016) and Blake et al. (2016a), respectively.

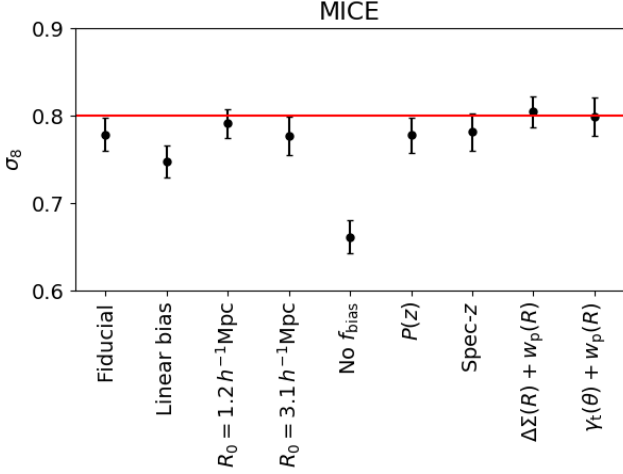


Fig. 10. Measurements of the σ_8 parameter marginalising over the galaxy bias factors, resulting from fits to the Υ_{gm} and Υ_{gg} statistics measured for the MICE lens redshift slices. We adopted a 2-parameter galaxy bias model ($b_{\text{lin}}, b_{\text{NL}}$) for each redshift slice, such that the fit varied 11 parameters in total. The left-hand point resulted from our fiducial analysis choice, assuming $R_0 = 2.0 h^{-1} \text{Mpc}$ and a point-based photo- z dilution correction. The second point from the left compares the result of a fit using only the linear galaxy bias parameter (setting $b_{\text{NL}} = 0$). We also show results varying the value of R_0 and the photo- z dilution correction. The horizontal red line indicates the MICE simulation fiducial cosmology $\sigma_8 = 0.8$.

We split the KiDS-1000 source catalogue into five different tomographic samples by the value of the BPZ photometric redshift, using the same bin divisions $z_B = [0.1, 0.3, 0.5, 0.7, 0.9, 1.2]$ adopted in Sect. 8. The effective source density of each tomographic sample is $n_{\text{eff}} = [0.88, 1.33, 2.04, 1.49, 1.26] \text{ arcmin}^{-2}$ (Hildebrandt et al. 2020), estimated using the method of Heymans et al. (2012). We divided the BOSS and 2dFLenS LRG catalogues into five spectroscopic redshift slices of width $\Delta z_1 = 0.1$ in the range $0.2 < z_1 < 0.7$.

We measured the average tangential shear $\gamma_t(\theta)$ and projected mass density $\Delta\Sigma(R)$ between all pairs of KiDS-1000 tomographic source samples and LRG redshift slices in the north and south regions, using the same estimators and binning as utilised for the MICE mocks in Sect. 8.1 and applying a multiplicative shear bias correction for each tomographic sample (Kannawadi et al. 2019). For the $\Delta\Sigma$ measurement, we again restricted the source-lens pairs such that $z_B > z_1$, and (in our fiducial analysis) applied a point-based photo- z dilution correction.

We generated an analytical covariance matrix for each measurement, initially using a fiducial lens linear bias factor $b_L = 2$, and iterating this value following a preliminary fit to the projected lens clustering. We tested that the analytical error determination agreed sufficiently well with a jack-knife error analysis where the regions were defined as the KiDS pointings; the results of this test and the overall analytical covariance are visually similar to their equivalents for the MICE mocks shown in Figs. 6 and 7, and we do not repeat these figures. We used the KV450 spectroscopic calibration sample with DIR weights (Hildebrandt et al. 2020) to estimate the redshift distribution of each tomographic source sample for use in the analytical covariance matrix, in the modelling of $\gamma_t(\theta)$ and in the distribution-based correction to $\Delta\Sigma(R)$ for photo- z dilution, and to determine the f_{bias} values for the point-based photo- z dilution correction.

We propagated the uncertainties in the multiplicative correction factors due to the shear calibration bias and photometric redshift dilution using the method described in Sect. 5.6. Regarding the multiplicative shear calibration, we followed Hildebrandt et al. (2020) in adopting an error $\sigma_m = 0.02$ that is fully correlated across all samples, such that $\text{Cov}[\alpha_{ij}, \alpha_{lm}] = \sigma_m^2$ in Eq. 66.

The sample variance in the spectroscopic training set can be characterised by an uncertainty in mean spectroscopic redshift which varies for each tomographic sample in the range $\sigma_z = 0.011 \rightarrow 0.039$ (see Hildebrandt et al. 2020, Table 2). We propagated these errors into the determination of f_{bias} by re-evaluating Eq. 35 shifting all the spectroscopic redshifts by a small amount to determine the derivatives $\partial f_{\text{bias},ij} / \partial z_j$, where i denotes the lens sample and j the source sample. Using error propagation, we then scaled the derivatives by the errors $\sigma_{z,j}$ to find the covariance matrix of the uncertainties,

$$\text{Cov}[f_{\text{bias},ij}, f_{\text{bias},lm}] = \frac{\partial f_{\text{bias},ij}}{\partial z_j} \frac{\partial f_{\text{bias},lm}}{\partial z_m} \sigma_{z,j}^2 \delta_{jm}^K, \quad (74)$$

where the final Kronecker delta δ_{jm}^K indicates that these uncertainties are correlated for different lens samples corresponding to the same source sample, but uncorrelated between source samples (we refer the reader to Joachimi et al. (in prep.) for further investigation of this point). This uncertainty can be propagated into the analytical covariance matrix using Eqs. 65 and 66 with $\text{Cov}[\alpha_{ij}, \alpha_{lm}] = \text{Cov}[f_{\text{bias},ij}, f_{\text{bias},lm}]$.

We used the analytical covariance matrices to combine the separate KiDS-N and KiDS-S measurements into a single joint estimate of the galaxy-galaxy lensing statistics and associated covariance, which we utilised in the remainder of this study (we test the consistency of the individual BOSS and 2dFLenS results in Sect. 9.4). We display the KiDS $\gamma_t(\theta)$ and $\Delta\Sigma(R)$ galaxy-galaxy lensing measurements in the different tomographic combinations in Figs. 3 and 4. We note again that there are some differences between the galaxy-galaxy lensing signals measured in the mocks and data, given that the mocks have not been tuned to reproduce the BOSS and 2dFLenS clustering properties. These differences are particularly evident on the smallest scales, owing to an inconsistent halo occupation. Our study does not require the mocks to precisely replicate the data in order to test our analysis framework.

We obtained the most accurate measurement of the projected correlation function $w_p(R)$ of each lens redshift slice using the full BOSS DR12 dataset, combining the LOWZ and CMASS selections and spanning 9376 deg^2 (Reid et al. 2016). We adopted the same spatial separation bins as for the MICE mocks, again assuming $\Pi_{\text{max}} = 100 h^{-1} \text{Mpc}$. When analysing the BOSS sample we included completeness weights but excluded ‘FKP’ weights (Feldman et al. 1994), which are designed to optimise the clustering signal-to-noise ratio but may not be appropriate in the case of galaxy-galaxy lensing. The full 2dFLenS dataset is too small to offer a competitive measurement of $w_p(R)$, although given that it was selected using BOSS-inspired colour-magnitude cuts, we assumed that the BOSS clustering is representative of the combined LRG sample (and we test this approximation in Sect. 9.4). Since the overlap of the KiDS-N source catalogue and full BOSS sample is small (4% of BOSS), we also assumed that the galaxy-galaxy lensing and clustering measurements are uncorrelated.

We combined the correlated $\Delta\Sigma$ measurements for each lens redshift slice, averaging over the different source samples. Fig. 11 displays these measurements, together with the projected

clustering $w_p(R)$ of the full BOSS sample, the corresponding $\Upsilon_{\text{gm}}(R, R_0)$ and $\Upsilon_{\text{gg}}(R, R_0)$ statistics assuming a fiducial choice $R_0 = 2 h^{-1}$ Mpc (we consider the impact of varying this choice in Sect. 9.4), and the direct $E_G(R)$ estimate using Eq. 44.

We generated fiducial cosmological models for these statistics using a non-linear matter power spectrum corresponding to the best-fitting ‘TTTEEE+lowE+lensing’ *Planck* cosmological parameters (Planck Collaboration et al. 2018). When producing the models overplotted in Figs. 3, 4 and 11, we determined best-fitting linear and non-linear galaxy bias parameters by fitting to the Υ_{gg} measurements for each lens redshift slice for the separation range $R > 5 h^{-1}$ Mpc, and applied these same bias parameters to the galaxy-galaxy lensing models. Values of the χ^2 statistic for each statistic and lens redshift slice, produced using these models and the analytical covariance, are displayed in each panel of Fig. 11, and indicate that the measurements are consistent with the model.

9.2. Redshift-space distortion inputs

We adopted values of the redshift-space distortion parameters β for the BOSS sample as a function of redshift by interpolating the literature analysis of Zheng et al. (2019), who provide RSD measurements in narrow redshift slices (which also agree with the compilation of results in Alam et al. 2017a). In order to interpolate these measurements to our redshift locations, which slightly differ from the bin centres of Zheng et al. (2019), we created a Gaussian process model for $\beta(z)$ and its errors, which lie in the range 12 – 20% as a function of redshift, using the sum of a Matern kernel and white noise kernel. We note that the error in β makes up roughly half the variance budget for E_G in the lowest lens redshift slice (i.e. increases the total error by $\sim \sqrt{2}$), but is subdominant for the other redshift slices.

9.3. Amplitude-ratio test E_G

We used the KiDS-1000 and LRG clustering and galaxy-galaxy lensing measurements, with the previously-published values of β , to determine a scale-independent value of the amplitude ratio, $\langle E_G \rangle$. We adopted the same fiducial analysis method as for the MICE mocks: we performed a joint fit to the Υ_{gm} and Υ_{gg} measurements for each redshift slice, varying A_E and the bias parameters b_L and b_{NL} as in Eq. 71, treating A_E as an additional amplitude parameter for Υ_{gm} . We then deduced $\langle E_G \rangle = A_E/\beta$, propagating the errors in β assuming all these statistics are independent (please see Sect. 4.3 for a note on this approximation). We used the analytical covariance matrices for these statistics, assumed $R_0 = 2 h^{-1}$ Mpc, and fitted the model to $R > 5 h^{-1}$ Mpc for Υ_{gg} and to all scales for Υ_{gm} . We consider the effect of varying these analysis choices in Sect. 9.4. In particular, we note that fitting the directly-determined $E_G(R)$ values (shown in the fifth row of Fig. 11) produced results which were entirely consistent with our fiducial analysis.

Our resulting fits for $\langle E_G \rangle$ were $[0.43 \pm 0.09, 0.45 \pm 0.07, 0.33 \pm 0.06, 0.38 \pm 0.07, 0.34 \pm 0.08]$ for redshifts $z = [0.25, 0.35, 0.45, 0.55, 0.65]$. The measurements have a small degree of correlation, owing to sharing a common source sample, and the analytical covariance matrix is listed in Table 1. We plot these measurements in Fig. 12, together with a literature compilation (Reyes et al. 2010; Blake et al. 2016b; Pullen et al. 2016; Alam et al. 2017b; de la Torre et al. 2017; Amon et al. 2018; Singh et al. 2019; Jullo et al. 2019). The thickness of the purple shaded stripe in Figs. 12 and 13 illustrates the 68% confidence

Table 1. The covariance matrix corresponding to our measurements $E_G = [0.43, 0.45, 0.33, 0.38, 0.34]$ at $z = [0.25, 0.35, 0.45, 0.55, 0.65]$. Each entry has been multiplied by 10^4 for clarity of display. The cross-correlation between different redshift slices is small.

Redshift	0.25	0.35	0.45	0.55	0.65
0.25	76.821	3.477	3.251	3.330	3.178
0.35	3.477	52.101	1.559	1.633	1.509
0.45	3.251	1.559	34.674	1.088	1.017
0.55	3.330	1.633	1.088	43.179	0.893
0.65	3.178	1.509	1.017	0.893	67.585

range of the prediction of the *Planck* ‘TTTEEE+lowE+lensing’ parameter chain at each redshift, assuming a flat Λ CDM Universe.

Our measurements provide the best existing determination of the lensing-clustering amplitude ratio (noting that previous measurements displayed in Figure 12 typically correspond to significantly wider lens redshift ranges than our study), which is consistent with matter density values $\Omega_m \sim 0.3$. Varying the Ω_m parameter within a flat Λ CDM cosmological model, assuming $E_G = \Omega_m/f = \Omega_m(0)/\Omega_m(z)^{0.55}$, we find $\Omega_m = 0.27 \pm 0.04$ (with a minimum $\chi^2 = 1.2$ for four d.o.f.). In principle in linear theory, this measurement is insensitive to the other cosmological parameters in a flat Λ CDM scenario. The resulting error in Ω_m is, naturally, somewhat larger than that provided by analyses utilising the full shape of the cosmic shear and clustering functions (e.g. Tröster et al. 2020), albeit requiring fewer model assumptions.

We tested for the scale dependence of the $E_G(R)$ measurements (in the fifth row of Fig. 11) by jointly fitting an empirical six-parameter model $E_G(R, z_i) = A_i [1 + \alpha \log_{10}(R)]$ to all the redshift slices, where α quantifies the fractional variation in E_G per decade in projected scale R (in h^{-1} Mpc) and A_i is a free amplitude for each of the five lens redshift slices. We obtained a 68% confidence region $\alpha = 0.17 \pm 0.26$ (with a minimum $\chi^2 = 35.9$ for 34 d.o.f.), which is consistent with no scale dependence, as predicted in the standard gravity scenario. These fits are displayed in Fig. 13.

9.4. Systematics tests

We now consider the effect on our cosmological fits of varying our fiducial analysis choices. Fig. 14 is a compilation of different determinations of $\langle E_G \rangle$ in each lens redshift slice. The upper-left panel compares the fits varying the photo- z dilution correction for $\Delta\Sigma$ between the point-based and distribution-based approaches, and the upper-right panel shows measurements varying the small-scale parameter R_0 . The lower-left panel considers separate determinations based on the BOSS and 2dFLenS samples (using the BOSS clustering measurements in both cases). The lower-right panel alters the fitting method to only use a linear-bias model (set $b_{\text{NL}} = 0$), and to use a direct fit to the $E_G(R)$ measurements presented in the fifth row of Fig. 11, as opposed to our fiducial fits to Υ_{gm} and Υ_{gg} . In all cases, the systematic variation of the recovered E_G values is negligible compared to the statistical errors (noting that the BOSS and 2dFLenS comparison is also subject to sample variance error). We find that varying the photo- z dilution correction, choice of R_0 and fitting method produce a systematic variation of $0.08\text{-}\sigma$, $0.24\text{-}\sigma$ and $0.38\text{-}\sigma$ respectively, when expressed as a fraction of the statistical error σ .

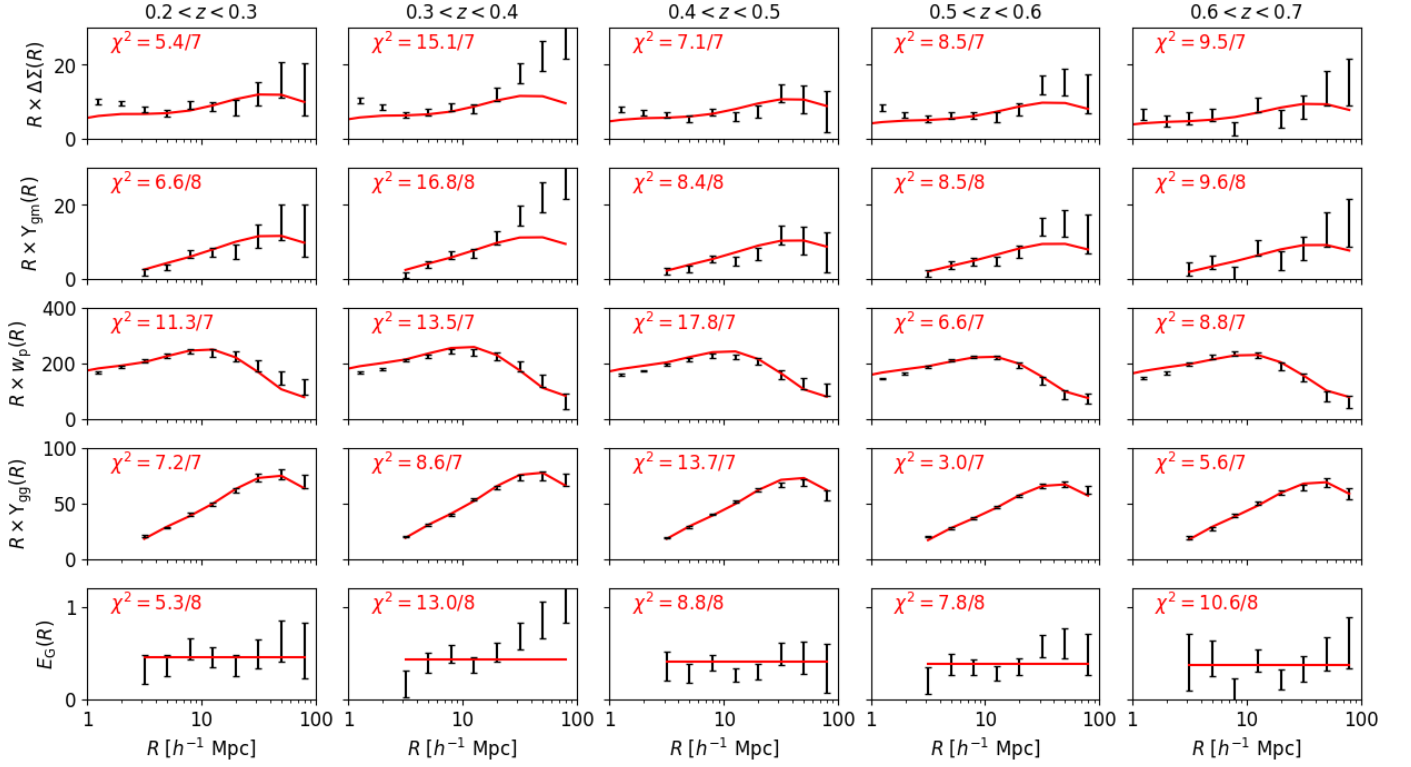


Fig. 11. Measurements of a series of galaxy-galaxy lensing and clustering statistics (rows) for each different LRG redshift slice (columns) correlated with the KiDS-1000 source sample, displayed in the same style as Fig. 5. We note that the first and third rows – $\Delta\Sigma(R)$ and $w_p(R)$ – represent the original galaxy-galaxy lensing and projected clustering measurements, and the remaining rows – $Y_{gm}(R)$, $Y_{gg}(R)$ and $E_G(R)$ – represent statistics derived from these original measurements.

10. Summary

We have used the latest weak gravitational lensing data release from the Kilo-Degree Survey, KiDS-1000, together with overlapping galaxy redshift survey data from BOSS and 2dFLenS, to perform cosmological tests associated with the relative amplitude of galaxy-galaxy lensing and galaxy clustering statistics. We quantified our results using the $E_G(R)$ statistic, which we were able to measure up to projected separations of $100 h^{-1}$ Mpc, recovering a scale-independent value of $\langle E_G \rangle$ with accuracies in the range 15-20% in five lens redshift slices of width $\Delta z = 0.1$. The scale-dependence and redshift-dependence of these measurements are consistent with the theoretical expectation of general relativity in a Universe with matter density $\Omega_m \sim 0.3$. The measurements are consistent with a scale-independent model for E_G , and constrain allowed variation within 25% ($1-\sigma$) error per decade in projected scale. Fitting our E_G dataset with a flat Λ CDM model, we find $\Omega_m = 0.27 \pm 0.04$.

We demonstrated that our results are robust against different analysis methodologies. In particular, we showed that:

- Source photometric redshift errors cause a significant dilution of the inferred projected mass density $\Delta\Sigma$, by causing unlensed foreground sources to appear in the background of lenses. We demonstrated that this dilution may be corrected either by modifying the estimator for $\Delta\Sigma$ to include the redshift probability distribution of the sources, or by using a source spectroscopic calibration sample to compute the multiplicative bias, producing consistent results. When applied to the mock catalogue, these estimation methods recovered the $\Delta\Sigma$ measurement obtained using source spectroscopic redshifts.

- A Gaussian analytical covariance for the galaxy-galaxy lensing statistics, with suitable modifications for the survey selection function and small-scale noise terms, predicted errors which agreed within 20% with estimates from a jack-knife procedure and from the variation across mock realisations.
- Our amplitude-ratio test, based on the annular differential density statistics Y_{gm} and Y_{gg} , is insensitive to the small-scale parameter R_0 adopted in these statistics, producing consistent results for choices in the range $1 < R_0 < 3 h^{-1}$ Mpc. We also obtained consistent results analysing BOSS and 2dFLenS separately.

We performed an additional series of tests jointly fitting the overall amplitude σ_8 to the galaxy-galaxy lensing and clustering statistics in our mocks, marginalising over linear and non-linear bias parameters. For our scenario, the projected galaxy-galaxy lensing measurements produced a slightly more accurate determination of σ_8 than the angular statistics after matching of scales, although the results are fully consistent.

Our analysis sets the stage for upcoming, increasingly accurate, cosmological tests using amplitude-ratio statistics, as gravitational lensing and galaxy redshift samples continue to grow. These datasets will continue to offer rich possibilities for placing tighter constraints on allowed gravitational physics.

Acknowledgements. We thank the anonymous referee for useful comments on the paper. CB is grateful to Rossana Ruggeri, Alexie Leauthaud, Johannes Lange and Sukhdeep Singh for valuable discussions on galaxy-galaxy lensing measurements.

We acknowledge support from: the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO) through project number CE110001020 (CB, KG); the European Research Council under grant numbers 647112 (MA, BG, CH, TT), 770935 (AD, HH, JLvD, AHW) and 693024 (SJ); the Polish Ministry of Science and Higher Education through grant

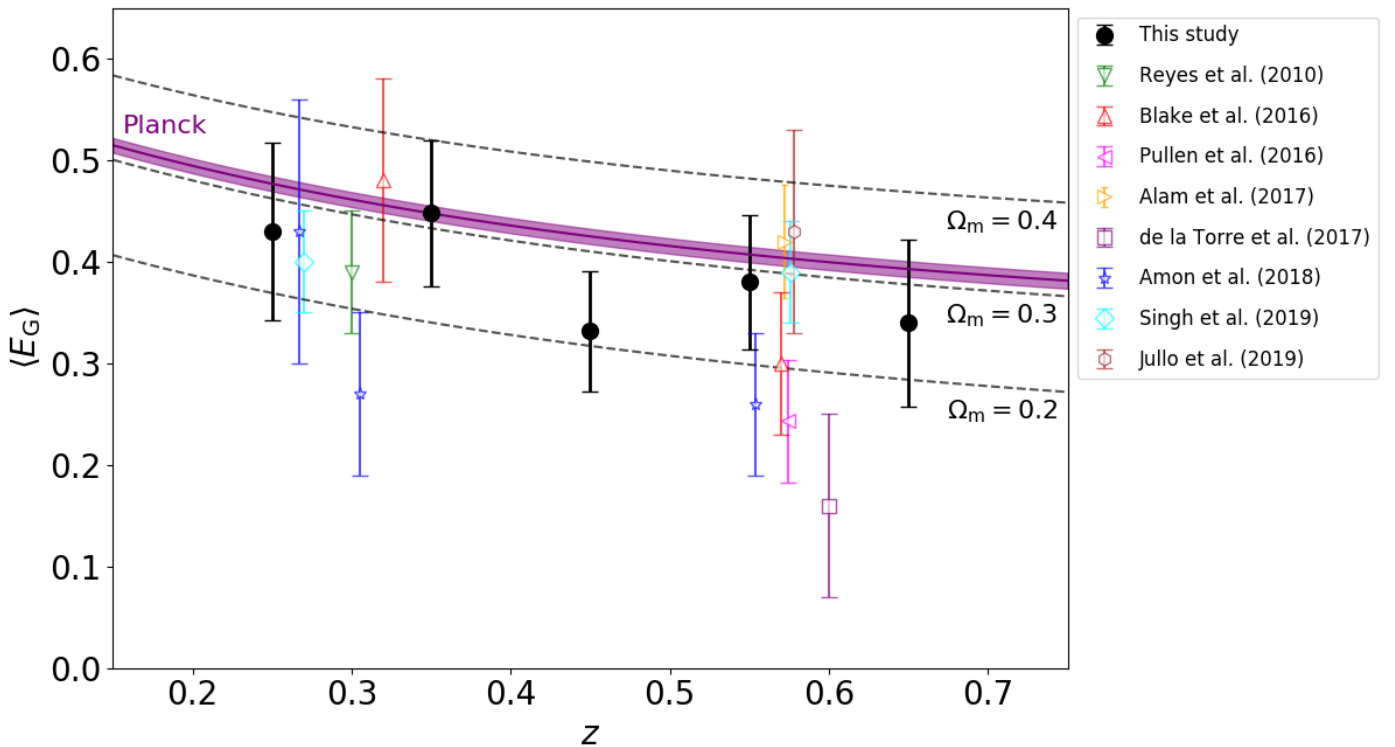


Fig. 12. Compilation of measurements of a scale-independent $\langle E_G \rangle$ as a function of redshift. The solid black data points illustrate the results of this study, in five redshift bins. These are compared to literature values from: Reyes et al. (2010) (green), Blake et al. (2016b) (red), Pullen et al. (2016) (magenta), Alam et al. (2017b) (orange), de la Torre et al. (2017) (purple), Amon et al. (2018) (blue), Singh et al. (2019) (cyan) and Jullo et al. (2019) (brown). We note that these previous measurements typically correspond to significantly wider lens redshift ranges than our study, and are also correlated in some cases owing to overlapping datasets. The thickness of the purple shaded stripe illustrates the 68% confidence range of the prediction of the *Planck* ‘TTTEEE+lowE+lensing’ parameter chain at each redshift assuming a flat Λ CDM Universe, and the dashed lines are the predictions of the model $E_G(z) = \Omega_m / f(z)$ for $\Omega_m = [0.2, 0.3, 0.4]$.

DIR/WK/2018/12 and the Polish National Science Center through grant 2018/30/E/ST9/00698 (MB); the Max Planck Society and the Alexander von Humboldt Foundation in the framework of the Max Planck-Humboldt Research Award endowed by the Federal Ministry of Education and Research (CH); Heisenberg grant Hi 1495/5-1 of the Deutsche Forschungsgemeinschaft (HH); the Beecroft Trust (SJ); Vici grant 639.043.512 financed by the Netherlands Organisation for Scientific Research (AK); the Alexander von Humboldt Foundation (KK); the NSFC of China grant 11973070, the Shanghai Committee of Science and Technology grant 19ZR1466600 and Key Research Program of Frontier Sciences grant ZDBS-LY-7013 (HYS); and the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant 797794 (TT).

The 2dFLenS survey is based on data acquired through the Australian Astronomical Observatory, under program A/2014B/008. It would not have been possible without the dedicated work of the staff of the AAO in the development and support of the 2dF-AAOmega system, and the running of the AAT.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

We have used `matplotlib` (Hunter 2007) for the generation of scientific plots, and this research also made use of `astropy`, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013).

Author contributions: All authors contributed to the development and writing of this paper. The authorship list is given in two groups: the lead author (CB), followed by an alphabetical group who have made a significant contribution to either the data products or to the scientific analysis.

References

- Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2018, *Phys. Rev. D*, 98, 043526
- Alam, S., Ata, M., Bailey, S., et al. 2017a, *MNRAS*, 470, 2617
- Alam, S., Miyatake, H., More, S., Ho, S., & Mandelbaum, R. 2017b, *MNRAS*, 465, 4853
- Amon, A., Blake, C., Heymans, C., et al. 2018, *MNRAS*, 479, 3422
- Asgari, M., Friswell, I., Yoon, M., et al. 2020a, arXiv e-prints, arXiv:2004.07811
- Asgari, M., Tröster, T., Heymans, C., et al. 2020b, *A&A*, 634, A127
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Baldauf, T., Smith, R. E., Seljak, U., & Mandelbaum, R. 2010, *Phys. Rev. D*, 81, 063531
- Bartelmann, M. & Schneider, P. 2001, *Phys. Rep.*, 340, 291
- Benítez, N. 2000, *ApJ*, 536, 571
- Bernstein, G. M. 2009, *ApJ*, 695, 652
- Bertschinger, E. 2011, *Philosophical Transactions of the Royal Society of London Series A*, 369, 4947
- Beutler, F., Seo, H.-J., Saito, S., et al. 2017, *MNRAS*, 466, 2242
- Blake, C., Amon, A., Childress, M., et al. 2016a, *MNRAS*, 462, 4240
- Blake, C., Joudaki, S., Heymans, C., et al. 2016b, *MNRAS*, 456, 2806
- Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, *Phys. Rep.*, 513, 1
- Crocce, M., Castander, F. J., Gaztañaga, E., Fosalba, P., & Carretero, J. 2015, *MNRAS*, 453, 1513
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
- de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, *A&A*, 582, A62

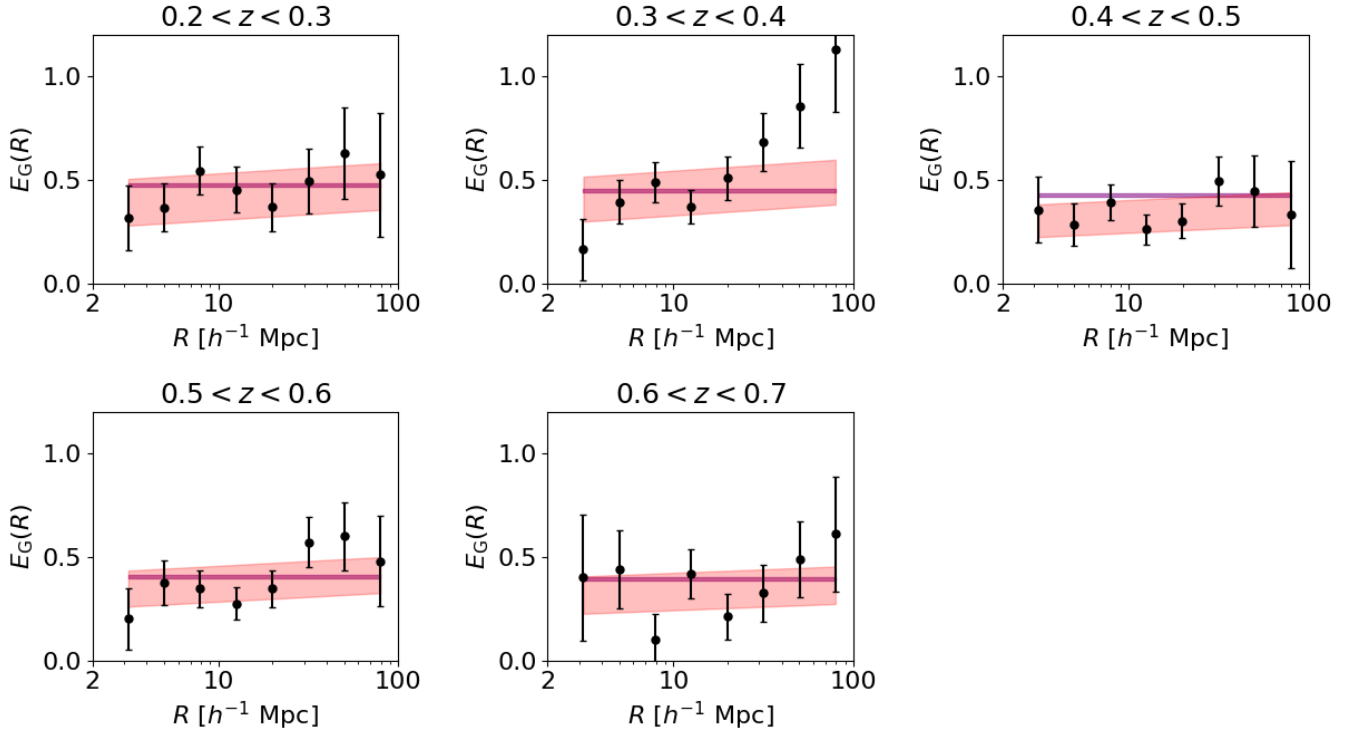


Fig. 13. Measurements of the $E_G(R)$ statistic for the KiDS-1000 source sample in each LRG lens redshift slice. Errors are derived from the diagonal elements of the analytical covariance matrix, propagating errors where appropriate. The thickness of the purple horizontal stripes illustrates the 68% confidence range of the prediction of the *Planck* ‘TTTEEE+lowE+lensing’ parameter chain at each redshift assuming a flat Λ CDM Universe, and the red bands illustrate the 68% confidence ranges of a scale-dependent model $E_G(R, z_i) = A_i [1 + \alpha \log_{10}(R)]$ discussed in Sect. 9.3.

de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, 175, 3
 de la Torre, S., Jullo, E., Giocoli, C., et al. 2017, *A&A*, 608, A44
 DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036
 Dvornik, A., Hoekstra, H., Kuijken, K., et al. 2018, *MNRAS*, 479, 1240
 Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, 142, 72
 Erben, T., Hildebrandt, H., Miller, L., et al. 2013, *MNRAS*, 433, 2545
 Euclid Collaboration, Blanchard, A., Camera, S., et al. 2019, arXiv e-prints, arXiv:1910.09273
 Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *ApJ*, 426, 23
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Fosalba, P., Crocce, M., Gaztañaga, E., & Castander, F. J. 2015a, *MNRAS*, 448, 2987
 Fosalba, P., Gaztañaga, E., Castander, F. J., & Crocce, M. 2015b, *MNRAS*, 447, 1319
 Guzik, J. & Seljak, U. 2001, *MNRAS*, 321, 439
 Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, *Nature*, 451, 541
 Hamilton, A. J. S. 1998, *Astrophysics and Space Science Library*, Vol. 231, *Linear Redshift Distortions: a Review*, ed. D. Hamilton, 185
 Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, *MNRAS*, 427, 146
 Hikage, C., Oguri, M., Hamana, T., et al. 2019, *PASJ*, 71, 43
 Hildebrandt, H., Köhlinger, F., van den Bosch, J. L., et al. 2020, *A&A*, 633, A69
 Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, *MNRAS*, 465, 1454
 Hu, W. & Jain, B. 2004, *Phys. Rev. D*, 70, 043009
 Hunter, J. D. 2007, *Computing In Science & Engineering*, 9, 90
 Huterer, D., Kirby, D., Bean, R., et al. 2015, *Astroparticle Physics*, 63, 23
 Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111
 Jain, B. & Khoury, J. 2010, *Annals of Physics*, 325, 1479
 Joachimi, B. & Bridle, S. L. 2010, *A&A*, 523, A1
 Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, *Space Sci. Rev.*, 193, 1
 Johnson, A., Blake, C., Amon, A., et al. 2017, *MNRAS*, 465, 4118
 Joudaki, S., Blake, C., Heymans, C., et al. 2017, *MNRAS*, 465, 2033
 Joudaki, S., Blake, C., Johnson, A., et al. 2018, *MNRAS*, 474, 4894
 Jullo, E., de la Torre, S., Cousinou, M. C., et al. 2019, *A&A*, 627, A137
 Kannawadi, A., Hoekstra, H., Miller, L., et al. 2019, *A&A*, 624, A92
 Kilbinger, M. 2015, *Reports on Progress in Physics*, 78, 086901
 Krause, E. & Eifler, T. 2017, *MNRAS*, 470, 2100
 Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, *A&A*, 625, A2

Landy, S. D. & Szalay, A. S. 1993, *ApJ*, 412, 64
 Lange, J. U., Yang, X., Guo, H., Luo, W., & van den Bosch, F. C. 2019, *MNRAS*, 488, 5771
 Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193
 Leauthaud, A., Saito, S., Hilbert, S., et al. 2017, *MNRAS*, 467, 3024
 Leonard, C. D., Ferreira, P. G., & Heymans, C. 2015, *J. Cosmology Astropart. Phys.*, 2015, 051
 Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
 Limber, D. N. 1953, *ApJ*, 117, 134
 Linder, E. V. 2005, *Phys. Rev. D*, 72, 043529
 Mandelbaum, R. 2018, *ARA&A*, 56, 393
 Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, *MNRAS*, 367, 611
 Mandelbaum, R., Slosar, A., Baldauf, T., et al. 2013, *MNRAS*, 432, 1544
 McDonald, P. 2006, *Phys. Rev. D*, 74, 103512
 McEwen, J. E., Fang, X., Hirata, C. M., & Blazek, J. A. 2016, *J. Cosmology Astropart. Phys.*, 2016, 015
 Miller, L., Heymans, C., Kitching, T. D., et al. 2013, *MNRAS*, 429, 2858
 Miyatake, H., Battaglia, N., Hilton, M., et al. 2019, *ApJ*, 875, 63
 Miyatake, H., More, S., Mandelbaum, R., et al. 2015, *ApJ*, 806, 1
 Nakajima, R., Mandelbaum, R., Seljak, U., et al. 2012, *MNRAS*, 420, 3240
 Park, Y., Rozo, E., & Krause, E. 2020, arXiv e-prints, arXiv:2004.07504
 Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv:1807.06209
 Pullen, A. R., Alam, S., He, S., & Ho, S. 2016, *MNRAS*, 460, 4098
 Reid, B., Ho, S., Padmanabhan, N., et al. 2016, *MNRAS*, 455, 1553
 Reyes, R., Mandelbaum, R., Seljak, U., et al. 2010, *Nature*, 464, 256
 Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
 Seitz, C. & Schneider, P. 1997, *A&A*, 318, 687
 Sheldon, E. S., Johnston, D. E., Frieman, J. A., et al. 2004, *AJ*, 127, 2544
 Shirasaki, M. & Takada, M. 2018, *MNRAS*, 478, 4277
 Singh, S., Alam, S., Mandelbaum, R., et al. 2019, *MNRAS*, 482, 785
 Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., & Vazquez Gonzalez, J. 2017, *MNRAS*, 471, 3827
 Smith, R. E., Hernández-Monteagudo, C., & Seljak, U. 2009, *Phys. Rev. D*, 80, 063528
 Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *MNRAS*, 341, 1311
 Song, Y.-S. & Percival, W. J. 2009, *J. Cosmology Astropart. Phys.*, 2009, 004

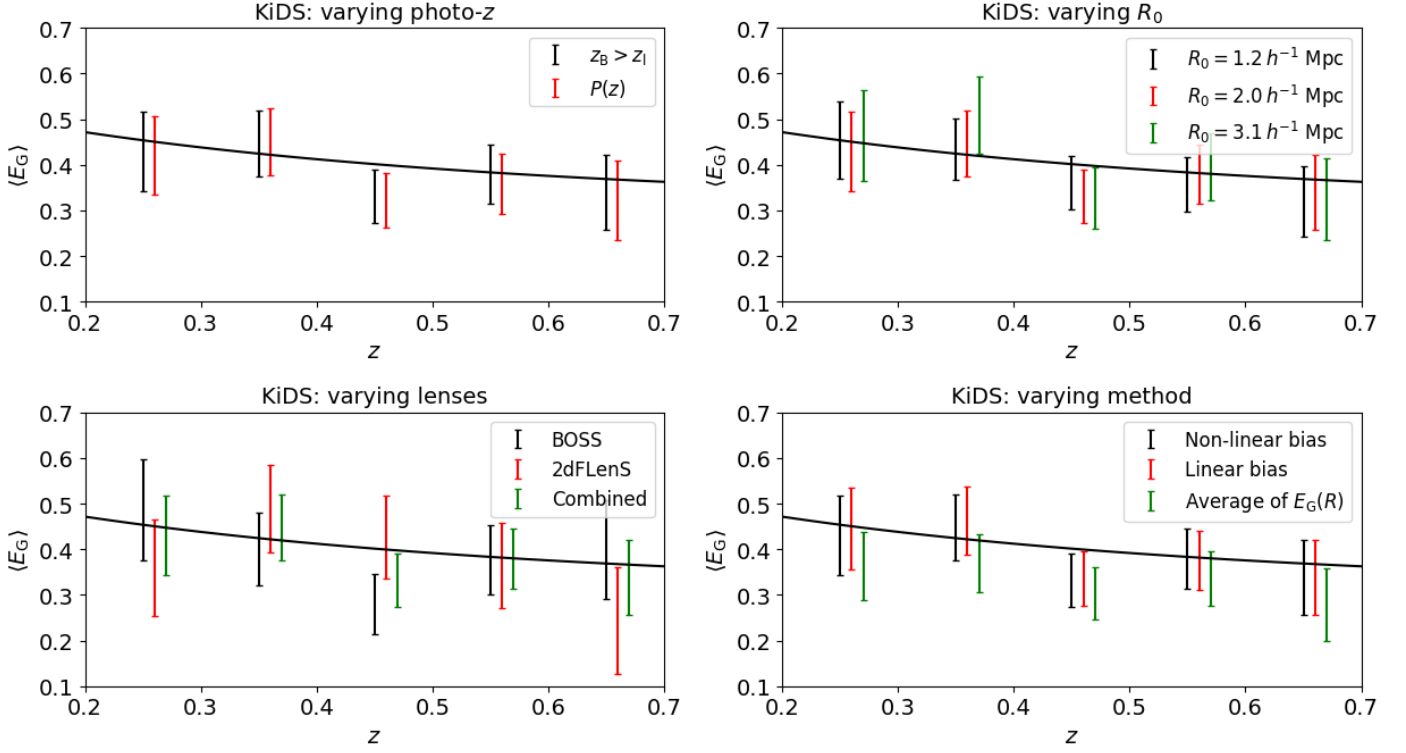


Fig. 14. Comparison of the scale-independent values $\langle E_G \rangle$ determined for each LRG redshift slice correlated with the KiDS-1000 source sample, varying the galaxy-galaxy lensing analysis assumptions, sample and methodology. Our fiducial analysis adopted a point-based photo- z correction with $z_B > z_l$, $R_0 = 2.0 h^{-1} \text{ Mpc}$ and a model fitted to $(\gamma_{gm}, \gamma_{gg})$ including non-linear galaxy bias. The upper-left panel compares determinations of $\langle E_G \rangle$ varying the photo- z dilution correction method, the upper-right panel varies the small-scale parameter R_0 , the lower-left panel varies the lens sample to consider BOSS and 2dFLenS separately, and the lower-right panel alters the fitting method to only include linear galaxy bias, and to use a direct fit to the $E_G(R)$ values. The model line in each case is the prediction $E_G(z) = \Omega_m / f(z)$, where $\Omega_m = 0.3$.

- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *ApJ*, 761, 152
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22
- Tröster, T., Sánchez, A. G., Asgari, M., et al. 2020, *A&A*, 633, L10
- Unruh, S., Schneider, P., Hilbert, S., et al. 2019, *arXiv e-prints*, arXiv:1910.06400
- Uzan, J.-P. & Bernardeau, F. 2001, *Phys. Rev. D*, 64, 083004
- Wang, Y. 2008, *J. Cosmology Astropart. Phys.*, 2008, 021
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *Phys. Rep.*, 530, 87
- Zhang, P., Liguori, M., Bean, R., & Dodelson, S. 2007, *Phys. Rev. Lett.*, 99, 141302
- Zheng, J., Zhao, G.-B., Li, J., et al. 2019, *MNRAS*, 484, 442

Appendix A: Covariance of average tangential shear

We may evaluate the covariance of γ_t between scales θ and θ' using Eq. 17,

$$\text{Cov}[\gamma_t^{ij}(\theta), \gamma_t^{kl}(\theta')] = \int \frac{d^2\ell}{(2\pi)^2} \int \frac{d^2\ell'}{(2\pi)^2} \text{Cov}[C_{gk}^{ij}(\ell), C_{gk}^{kl}(\ell')] J_2(\ell\theta) J_2(\ell'\theta'), \quad (\text{A.1})$$

where γ_t^{ij} denotes the average tangential shear of source sample j around lens sample i . We adopt an approximation that different multipoles ℓ are uncorrelated such that,

$$\text{Cov}[C_{gk}^{ij}(\ell), C_{gk}^{kl}(\ell')] = \delta_D(\ell - \ell') \sigma^2(\ell), \quad (\text{A.2})$$

where δ_D is the Dirac delta function, and the variance $\sigma^2(\ell)$ is given by Eq. 46.

Eq. A.1 becomes, after integrating the delta function $\int d^2\ell' f(\ell') \delta_D(\ell - \ell') = \frac{(2\pi)^2}{\Omega} f(\ell)$ where Ω is the total survey angular area in steradians,

$$\text{Cov}[\gamma_t^{ij}(\theta), \gamma_t^{kl}(\theta')] = \frac{1}{\Omega} \int \frac{d\ell}{2\pi} \sigma^2(\ell) J_2(\ell\theta) J_2(\ell'\theta'). \quad (\text{A.3})$$

If the measurements of γ_t are averaged within angular bins m and n , where the angular area of the i th bin is Ω_i (i.e. the area of the annulus between the bin limits), then the covariance between the bins is,

$$\begin{aligned} C_{mn} &= \int_m \frac{d^2\theta}{\Omega_m} \int_n \frac{d^2\theta'}{\Omega_n} \text{Cov}[\gamma_t^{ij}(\theta), \gamma_t^{kl}(\theta')] \\ &= \frac{1}{\Omega} \int \frac{d\ell}{2\pi} \sigma^2(\ell) \overline{J_{2,m}(\ell)} \overline{J_{2,n}(\ell)}, \end{aligned} \quad (\text{A.4})$$

where $\overline{J_{2,n}(\ell)} = \int_{\theta_{1,n}}^{\theta_{2,n}} \frac{2\pi\theta d\theta}{\Omega_n} J_2(\ell\theta)$. We also note that the contribution of any constant term in the covariance $\sigma^2(\ell) = C$ (such as the noise terms) is,

$$\begin{aligned} C_{mn} &= C \int \frac{2\pi\theta d\theta}{\Omega_m} \int \frac{2\pi\theta' d\theta'}{\Omega_n} \frac{1}{\Omega} \int \frac{d\ell}{2\pi} J_2(\ell\theta) J_2(\ell\theta') \\ &= \frac{C}{\Omega \Omega_n} \delta_{mn}^K, \end{aligned} \quad (\text{A.5})$$

using the Bessel function relation $\int_0^\infty J_n(ax) J_n(bx) x dx = \delta_D(a-b)/b$.

Appendix B: Modification of covariance for survey window

We derive how the covariance of a cross-correlation function between two Gaussian fields, $\delta_1(\mathbf{x})$ and $\delta_2(\mathbf{x})$, is modified by the window function of the fields, $W_1(\mathbf{x})$ and $W_2(\mathbf{x})$. We adopt the case of a 2D flat sky, where the vector separation \mathbf{r} between two points has magnitude r and orientation angle θ . An estimator of the cross-correlation function of the fields at separation r is,

$$\hat{\xi}(r) = \frac{1}{A_2(r)} \int \frac{d\theta}{2\pi} \int d^2\mathbf{x} \delta_1(\mathbf{x}) \delta_2(\mathbf{x}+\mathbf{r}) W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r}), \quad (\text{B.1})$$

where $A_2(r) = \int \frac{d\theta}{2\pi} \int d^2\mathbf{x} W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r})$. The expectation value of this expression is,

$$\begin{aligned} \langle \hat{\xi} \rangle &= \frac{1}{A_2(r)} \int \frac{d\theta}{2\pi} \int d^2\mathbf{x} \langle \delta_1(\mathbf{x}) \delta_2(\mathbf{x}+\mathbf{r}) \rangle W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r}) \\ &= \frac{1}{A_2(r)} \int \frac{d^2\mathbf{k}}{(2\pi)^2} P_{12}(\mathbf{k}) \int \frac{d\theta}{2\pi} \int d^2\mathbf{x} W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \\ &\approx \frac{1}{A_2(r)} \frac{1}{2\pi} \int dk k P_{12}(k) \int \frac{d\theta}{2\pi} A_2(r) e^{-ikr \cos \theta} \\ &= \frac{1}{2\pi} \int dk k P_{12}(k) J_0(kr), \end{aligned} \quad (\text{B.2})$$

where we have introduced the cross-power spectrum $P_{12}(k)$, and the approximation in the third line of Eq. B.2 ignores the θ dependence of $\int d^2\mathbf{x} W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r})$.

The covariance of the estimator may be deduced from,

$$\begin{aligned} \langle \hat{\xi}(\mathbf{r}) \hat{\xi}(\mathbf{s}) \rangle &= \frac{1}{A_2(r) A_2(s)} \int d^2\mathbf{x} \int d^2\mathbf{y} \\ &A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{y}, \mathbf{s}) \langle \delta_1(\mathbf{x}) \delta_2(\mathbf{x}+\mathbf{r}) \delta_1(\mathbf{y}) \delta_2(\mathbf{y}+\mathbf{s}) \rangle, \end{aligned} \quad (\text{B.3})$$

where we have written $A_{12}(\mathbf{x}, \mathbf{r}) = W_1(\mathbf{x}) W_2(\mathbf{x}+\mathbf{r})$. Expanding this expression using Wick's theorem for a Gaussian random field, $\langle \delta_1 \delta_2 \delta_3 \delta_4 \rangle = \langle \delta_1 \delta_2 \rangle \langle \delta_3 \delta_4 \rangle + \langle \delta_1 \delta_3 \rangle \langle \delta_2 \delta_4 \rangle + \langle \delta_1 \delta_4 \rangle \langle \delta_2 \delta_3 \rangle$, we find

$$\begin{aligned} \text{Cov}[\hat{\xi}(\mathbf{r}), \hat{\xi}(\mathbf{s})] &= \frac{1}{A_2(r) A_2(s)} \int d^2\mathbf{x} \int d^2\mathbf{y} \\ &A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{y}, \mathbf{s}) [\langle \delta_1(\mathbf{x}) \delta_1(\mathbf{y}) \rangle \langle \delta_2(\mathbf{x}+\mathbf{r}) \delta_2(\mathbf{y}+\mathbf{s}) \rangle \\ &+ \langle \delta_1(\mathbf{x}) \delta_2(\mathbf{y}+\mathbf{s}) \rangle \langle \delta_2(\mathbf{x}+\mathbf{r}) \delta_1(\mathbf{y}) \rangle]. \end{aligned} \quad (\text{B.4})$$

Using $\langle \delta_i(\mathbf{x}) \delta_j(\mathbf{y}) \rangle = \langle \delta_i(\mathbf{x}) \delta_j^*(\mathbf{y}) \rangle = \int \frac{d^2\mathbf{k}}{(2\pi)^2} P_{ij}(\mathbf{k}) e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}$, and omitting some algebra, the first term evaluates to,

$$\int \frac{d^2\mathbf{k}}{(2\pi)^2} P_{11}(\mathbf{k}) P_{22}(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int d^2\mathbf{x} A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{x}, \mathbf{s}), \quad (\text{B.5})$$

and the second term evaluates to,

$$\int \frac{d^2\mathbf{k}}{(2\pi)^2} P_{12}^2(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int d^2\mathbf{x} A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{x}, \mathbf{s}). \quad (\text{B.6})$$

The expression for the covariance is then,

$$\begin{aligned} \text{Cov}[\hat{\xi}(\mathbf{r}), \hat{\xi}(\mathbf{s})] &= \frac{\int d^2\mathbf{x} A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{x}, \mathbf{s})}{A_2(r) A_2(s)} \\ &\int \frac{d^2\mathbf{k}}{(2\pi)^2} [P_{11}(\mathbf{k}) P_{22}(\mathbf{k}) + P_{12}^2(\mathbf{k})] e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})}. \end{aligned} \quad (\text{B.7})$$

Averaging the estimator over angles we obtain,

$$\begin{aligned} \text{Cov}[\hat{\xi}(r), \hat{\xi}(s)] &\approx \frac{A_3(r, s)}{A_2(r) A_2(s)} \\ &\frac{1}{2\pi} \int dk k [P_{11}(k) P_{22}(k) + P_{12}^2(k)] J_0(kr) J_0(ks), \end{aligned} \quad (\text{B.8})$$

where $A_3(r, s) = \int d^3\mathbf{r} \int d^3\mathbf{s} \int d^2\mathbf{x} A_{12}(\mathbf{x}, \mathbf{r}) A_{12}(\mathbf{x}, \mathbf{s})$.

Appendix C: Combining correlated tomographic slices

In order to reduce the size of a data vector, we can optimally combine separate correlated estimates of a statistic, such as a galaxy-galaxy lensing measurement for a given lens sample against different tomographic source slices. This procedure is an example of data compression (Tegmark et al. 1997).

Suppose we have measured a given statistic at N_r different scales, for N_s source tomographic slices, and we wish to average the statistic over source samples, where the measurements in the different slices are correlated. We'll arrange these quantities in a data vector \mathbf{x} of length $N_s N_r$ with corresponding covariance matrix \mathbf{C} of dimension $N_s N_r \times N_s N_r$. The operation to combine the different tomographic slices to a compressed data vector \mathbf{y} of length N_r can be written as,

$$\mathbf{y} = \mathbf{w}^T \mathbf{x}, \quad (\text{C.1})$$

where \mathbf{w} is a weight matrix of dimension $N_s N_r \times N_r$, and we normalise the weights such that the column corresponding to each scale bin sums to unity. The optimal choice of weight matrix (Tegmark et al. 1997) is,

$$\mathbf{w} = \mathbf{C}^{-1} \mathbf{D}, \quad (\text{C.2})$$

where \mathbf{D} is a matrix of dimension $N_s N_r \times N_r$, whose columns consist of $N_s N_r$ entries for each final scale bin, with value 1 when the entry in \mathbf{x} corresponds to the same scale bin, and value 0 otherwise. The resulting covariance matrix of \mathbf{y} is,

$$\mathbf{C}_y = \mathbf{w}^T \mathbf{C} \mathbf{w}, \quad (\text{C.3})$$

which has dimension $N_r \times N_r$.

We found that this data compression scheme is more robust against numerical issues with the matrix inverse (and suffers negligible loss in precision) if we replaced the weight matrix with $\mathbf{w} = \mathbf{V}^{-1} \mathbf{D}$, where \mathbf{V} is a diagonal matrix just containing the variance of the measurements. In this implementation the weight matrix is slightly suboptimal, but we retained the full covariance matrix in Eq. C.3 to ensure correct error propagation.